



POLITECNICO

MILANO 1863

- Dipartimento di Chimica, Materiali e Ingegneria Chimica “Giulio Natta”
- Center For Nano Science and Technology @ PoliMi

Molecular bases of Phosphodiesterase-4D inhibition by memory-enhancing GEBR-library compounds

Coordinatore: Prof. Alessio Frassoldati

Tutor: Prof. Maurizio Masi

Relatore: Dr. Emilio Parisini

Ciclo: 31°

*“In the beginning the Universe was created.
This has made a lot of people very angry
and been widely regarded as a bad move “*

The restaurant at the end of the Universe - Douglas Adams

Index

INDEX	2
ABSTRACT	5
NOTES AND ABBREVIATIONS:	8
MOLECULAR BASES OF PHOSPHODIESTERASE-4D INHIBITION BY MEMORY-ENHANCING GEBR-LIBRARY COMPOUNDS	10
1 INTRODUCTION	10
1.1 CELL SIGNALING	10
1.1.1 <i>Delivering a signal</i>	10
1.1.2 <i>Recognition</i>	11
1.1.3 <i>G protein-coupled receptors</i>	13
1.1.4 <i>Transmission</i>	14
1.2 THE CAMP PATHWAY	15
1.2.1 <i>Synthesis</i>	16
1.2.2 <i>Effectors</i>	16
1.3 CAMP-SPECIFIC PHOSPHODIESTERASE	18
1.3.1 <i>Coordination and catalysis</i>	18
1.3.2 <i>PDE4 gene structure</i>	19
1.3.3 <i>Dimerization</i>	20
1.3.4 <i>HARBS and LARBS</i>	23
1.3.5 <i>PDE4D Regulation by phosphorylation</i>	24
1.3.6 <i>PDE4D localization and interactions</i>	26
1.3.7 <i>The pocket</i>	28
1.4 PHOSPHODIESTERASE 4 AS PHARMACEUTICAL TARGET	31
1.4.1 <i>Side effects</i>	32
1.4.2 <i>Airway disease</i>	33
1.4.3 <i>Central Nervous System</i>	33
1.4.4 <i>Alzheimer's disease</i>	34
1.4.5 <i>X-fragile</i>	35
1.5 PHOSPHODIESTERASE 4 INHIBITORS	36
1.5.1 <i>Structural bases</i>	36
1.5.2 <i>Allosteric modulation</i>	40

1.6	THE GEBR LIBRARY	42
1.6.1	<i>Historical overview</i>	42
1.6.2	<i>Chemical bases of the library</i>	47
1.7	AIM AND OVERVIEW OF THE PhD PROJECT: A MECHANIST UNDERSTANDING OF PDE4D INHIBITION BY GEBR-LIBRARY COMPOUNDS	49
1.7.1	<i>Origins</i>	49
1.7.2	<i>Expanding the project</i>	49
2	MATERIALS AND METHODS	52
2.1	PHOSPHODIESTERASE 4D CATALYTIC DOMAIN EXPRESSION AND PURIFICATION	52
2.1.1	<i>Construct</i>	52
2.1.2	<i>Protein Expression</i>	54
2.1.3	<i>Purification</i>	55
2.2	PHOSPHODIESTERASE 4D3 EXPRESSION AND PURIFICATION	58
2.2.1	<i>Construct</i>	58
2.2.2	<i>Bacmid generation</i>	59
2.2.3	<i>Generation of the virus – V0</i>	62
2.2.4	<i>Amplification of the virus – V1,2</i>	62
2.2.5	<i>Test expression</i>	63
2.2.6	<i>Purification</i>	65
2.3	PDE4 CATALYTIC DOMAIN X-RAY CRYSTALLOGRAPHY	68
2.3.1	<i>Crystal growth - first screening</i>	68
2.3.2	<i>Crystal growth - optimization</i>	69
2.3.3	<i>Crystal soaking experiments</i>	69
2.3.4	<i>X-ray diffraction collection</i>	70
2.3.5	<i>Structure solution</i>	71
2.4	ENZYMATIC ASSAYS	72
2.4.1	<i>Sensor system</i>	72
2.4.2	<i>Reagents details</i>	73
2.4.3	<i>Experimental Setup</i>	74
2.4.4	<i>Data analysis</i>	75
2.5	MOLECULAR DYNAMICS SIMULATIONS	76
2.5.1	<i>Model optimization</i>	76
2.5.2	<i>Simulations</i>	77
3	RESULTS AND DISCUSSION	79
3.1	INSIDE THE GEBR7B-LIKE COMPOUNDS: ROLE OF THE LONGER CHAIN AND OF THE HYDROXYL MOIETY	88
3.2	ROLE OF THE FLUORINE ATOMS	90

3.3	GEBR-32A TAUTOMERISM	92
3.4	MOLECULAR DYNAMICS	93
3.5	PYRROLE COMPOUNDS	95
3.6	STRUCTURE-BASED DESIGN OF NOVEL COMPOUNDS	97
3.6.1	<i>Imidazole compounds</i>	97
3.6.2	<i>Condensed cycles</i>	97
3.6.3	<i>Tail tuning</i>	99
4	CONCLUSIONS	101
	STRUCTURAL INSIGHTS INTO THE INTERPLAY BETWEEN CIS- AND TRANS-INTERACTIONS IN CADHERIN-MEDIATED CELL ADHESION	103
5	INTRODUCTION	103
5.1	MECHANISM OF CADHERIN DIMERIZATION	104
5.2	THE INTERPLAY BETWEEN THE <i>TRANS</i> AND THE <i>CIS</i> INTERFACE.	107
5.3	CADHERINS AS “UNDRUGGABLE” TARGETS	108
5.4	AIM OF THE PROJECT	110
6	MATERIALS AND METHODS	112
6.1	CONSTRUCTS	112
6.2	PROTEIN EXPRESSION AND PURIFICATION	114
6.2.1	<i>Preliminary purification</i>	115
6.2.2	<i>Cadherin activation:</i>	117
6.2.3	<i>Separation of the activated cadherins</i>	117
6.3	X-RAY CRYSTALLOGRAPHY	118
7	RESULTS AND DISCUSSION	120
	REFERENCES	125
	ACKNOWLEDGMENTS	136

Abstract

The neurological mechanisms governing the learning and memory processes depend critically on the cerebral levels of the second messenger cyclic adenosine monophosphate (cAMP) and on the correct functioning of the cAMP/PKA/CREB pathway. In the brain, cAMP levels are regulated by the activity of the type 4 phosphodiesterase (PDE4), an enzyme that hydrolyzes cAMP to 5'-AMP. Consistently, PDE4 is considered an important pharmaceutical target due to its crucial involvement in the signaling of the central nervous system. Indeed, some PDE4 inhibitors (PDE4Is) developed over time have been shown to improve memory and cognitive functions both under physiological and pathological conditions.

PDE4Is have mostly been designed to interact with the catalytic pocket of the enzyme and compete with the cAMP hydrolysis process. As a result, albeit providing interesting pro-cognitive and anti-depressant properties, current standard inhibitors such as Rolipram show also severe side effects due to their lack of isoform-specific binding properties. As many as 20 PDE4 isoforms that are generated by alternative mRNA splicing have so far been identified, each of them having a specific pattern of expression. The long isoforms of the genes include two regulatory domains, UCR1 and UCR2, while shorter isoforms are characterized by the absence of UCR1 or by the absence of both UCR1 and a portion of UCR2. Moreover, further variability is introduced by isoform-specific N-terminal regions, which are responsible for the interactions with signaling partners. The most extensive structural studies demonstrated that UCR1 and UCR2 mediate enzyme dimerization and that both UCR2 and the CR3 (a third regulatory region) domains are involved in the capping of the catalytic pocket of the enzyme, thus partially blocking access of either cAMP or PDE4 inhibitors. As a result, long dimeric isoforms, and short monomeric ones feature substantially different enzymatic and pharmacologic properties.

Burgin et al. demonstrated that it is possible to design allosteric modulators that are partially specific for PDE4D and that are capable of stabilizing the closed conformation of the UCR2 domain of PDE4D over the catalytic pocket through the clamping of a central phenylalanine. Likewise, a similar concept has been suggested also for the design of PDE4B inhibitors that stabilize the capping of the catalytic domain by CR3.

Currently, the most promising strategy to obtain a suitable therapeutic window while avoiding side effects seems to be the design of PDE4D selective inhibitors. Indeed, PDE4D has been demonstrated to be particularly important in memory development and its association with cognition processes has been demonstrated in knock-out mice, which have shown memory improvements similar to those shown by mice treated with Rolipram. Moreover, activating missense mutations affecting PDE4D activity were recently discovered to be associated with acrodysostosis, a rare genetic disease that brings also a phenotype of mental retardation.

In an effort to develop a drug that selectively targets the PDE4D isoforms without side effects, a number of compounds, commonly referred to as the GEBR library, that show partial selectivity for the PDE4D isoform have been synthesized. The reference inhibitors of the GEBR library are GEBR-7b and GEBR-32a, which have been shown to improve spatial and objects recognition memory and to increase hippocampal levels of cAMP in transgenic mice. By addressing the biochemical behavior of these compounds both at a structural and functional level, we aim at identifying key structural features that may provide the molecules with the ability to bind stably into the catalytic pocket while selectively interacting with the regulatory domains of the enzyme. The relatively high number of compounds synthesized to date (up to 90), some of which have interesting pharmacological and toxicological profiles, have so far provided useful, albeit not definitive, information about the inhibitors-enzyme interaction. Therefore, we set out to investigate the structure-function relationship for some selected GEBR compounds and contribute to the elucidation of the molecular bases of PDE4 inhibition.

Our characterization identified interesting structural properties, as well as three major families of compounds that differently dock inside PDE4D catalytic pocket with different conformations: extended, twisted and protruding. Among them, the protruding compounds feature a tail that develops within the S-pocket, pointing towards the external part of the pocket, where a putative interaction with the regulatory UCR2 helix could in principle be exploited. Every compound has been functionally tested *in-vitro* both on the PDE4D3 long isoform as well as on the PDE4D catalytic domain only, in order to assess whether a differential potency towards the two systems is present. Moreover, the structure of the PDE4D catalytic domain in complex with GEBR-7b and GEBR-32a were used in combination with the most extended X-ray crystal structure of a PDE4B

long isoform, in order to investigate the properties of our compounds in a fully-regulated system by means of molecular dynamics simulations.

The combination of the structural, functional and *in-silico* information derived from the study allowed us to design the new generation of GEBR molecules there are currently being synthesized in the Bruno's laboratory at the University of Genova.

Cadherins are a family of cell-adhesion transmembrane proteins that generate an adhesive surface by clustering on the cell surface and undergoing a homo-dimerization process with their partners on an adjacent cell. From an energetic point of view, the molecular events that take place during this multi-step recognition and dimerization process occur over a shallow potential energy surface that defines a complex pathway featuring multiple metastable intermediates. A deep understanding of the cadherin-mediated cell adhesion process is crucial not only because it is at the basis of numerous cell functions such as growth, differentiation and signaling, but also because cadherins have been identified as possible pharmaceutical targets for cancer and for inflammatory diseases. Indeed, the adhesive phenotype of tumor cells is usually perturbed relative to healthy cells, most often because of an aberrant cadherin expression profile.

In the Parisini's group, the molecular bases of cadherin activation and cadherin inhibition has long been a major area of research involving all the lab members. In this thesis, I present the structural characterization of two P-cadherin mutants (A5P P123G and A5P K14E P123G) in the framework of an ongoing study that aims at the comprehension of the relationship between two of the most important stabilization contacts along the cadherin activation trajectory: the so-called *cis*- and *trans*- interactions.

Notes and abbreviations:

- Chapter 3 contains extracts from the article: *Prodocimi T., Mollica L., Donini S., Semrau MS., Lucarelli AP., Aiolfi E., Cavalli A., Storici P., Alfei S., Brullo C., Bruno O., Parisini E. Molecular Bases of PDE4D Inhibition by Memory-Enhancing GEBR Library Compounds. Biochemistry 57, 2876–2888 (2018)*. The article has been published in the context of my PhD research and I contributed to the design and execution of all the experiments (gene design and cloning, protein production and purification, X-ray crystallography, biochemical assays, in silico simulations) as well as to writing of the manuscript.
- Throughout this thesis, the numbering scheme adopted for the residues of the different phosphodiesterase isoforms is relative to the sequence of phosphodiesterase 4D3, if not specified otherwise
- PDE: Phosphodiesterase
- PKA: Protein Kinase A
- AC: Adenylyl Cyclases
- AKAP: A Kinase Anchoring Proteins
- ERK: Extracellular-signal Related Kinases
- AD: Alzheimer's Disease
- COPD: Cronic Obstructive Pulmonary Disease
- FXS: Fragile-X Syndrome
- HARBS/LARBS: High/Low Affinity Rolipram Binding Site
- UCR: Unique Control Region
- CR: Control Region
- LR: Linker Region
- NAM: Negative Allosteric Modulator
- OLT: Object Location Task
- LTP: Long Term Potentiation
- PCR: Polymerase Chain Reaction

Molecular bases of Phosphodiesterase-4D inhibition by memory-enhancing GEBR-library compounds

1 Introduction

1.1 Cell signaling

By cell signaling we generally refer to the series of communication mechanisms and events that govern cellular activity, allowing cells to correctly respond to their microenvironment. This ability is provided by a complex network of molecular pathways that receive, transmit and process information coming from another cell compartment, from the external environment, from nearby cells and from distant cells.

Signal transduction is the process that allows a chemical or physical signal to be transmitted through a cell by means of a series of molecular events. This type of mechanism starts from a stimulus detected by a protein (receptor or sensor) that, in turn, activates a signaling cascade. The first actor of this process, for instance a neurotransmitter, is called “first messenger”; the receptor that senses the stimulus is called “signal transducer” and the proteins that will be subsequently activated are called “primary, secondary, etc. effectors”. Hence, the signaling pathway is classified relative to a specific stimulus.^{1,2}

1.1.1 Delivering a signal

Within cells, signals are always transduced along biochemical pathways. However, nature has adopted different ways to develop cellular communication routes¹:

- Extracellular messengers: signals are delivered by molecules that interact with partners on the target cells. First messengers can reach their target either by simple diffusion or by circulation, the latter usually occurring for long-range signaling and requiring a carrier protein to transport the first messenger.
- Gap junctions: communication between neighboring cells via channels that allow the direct exchange of signaling molecules.
- Interactions between surface proteins: signal is exchanged upon formation of an interaction between two proteins, one from each partner cell. The signal is then transduced in the form of a biochemical cascade.
- Electrical signaling: changes in the membrane potentials can be used to carry an information across the biological barrier.

1.1.2 Recognition

The next event in the signaling process consist in the very first step of the signal transduction. Target cells must be able to register a specific stimulus and to transduce it. Indeed, specialized proteins (receptors) are able to recognize the messenger substance or to respond to a physical stimulus, transmitting the signal to the primary effectors. Incoming signals can either be received by¹:

- Membrane receptors: in this case the signal is registered outside the cell and transduced across the membrane. There is no need for the first messenger to enter in the cell. Membrane receptors constitute the most abundant class of receptors. They feature a modular architecture with an extracellular domain, a transmembrane domain and an intracellular domain. As the first messenger interacts with the extracellular domain of the receptor, it produces a modification that triggers a change in the intracellular domain, which in turn allows the signal cascade to proceed further.

- Intracellular receptors: in this case, it is necessary for the first messenger to penetrate in the cell. Intracellular receptors are normally localized in the nucleus or in the cytoplasm and they function as signaling effectors upon binding to the first messenger.

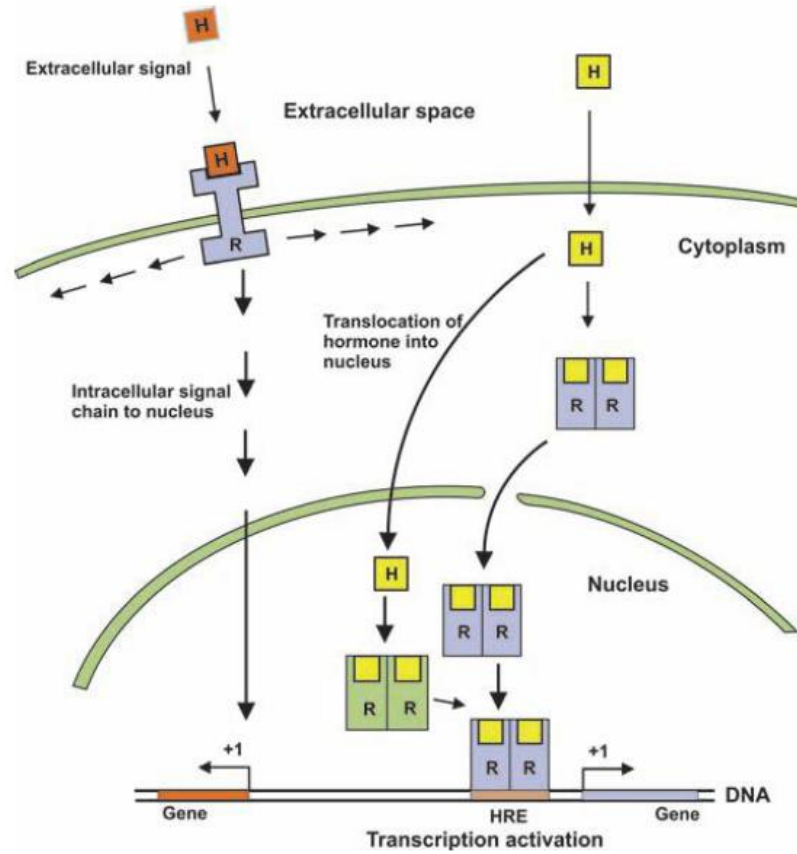


Figure 1
 Krauss, G. Biochemistry of Signal Transduction and Regulation
 Transmembrane and intracellular receptors.

1.1.3 G protein-coupled receptors

Among all the membrane receptor families, the largest is represented by the G protein-coupled receptors (GPCRs). Indeed, its almost 800 members are able to detect different extracellular signals and trigger a broad range of signaling cascades. Upon binding to their primary messengers, GPCRs transmit a signal by means of a conformational shift that allows other signaling proteins to interact with their cytosolic domain. It is important to note that thinking of GPCRs as on/off switches of a

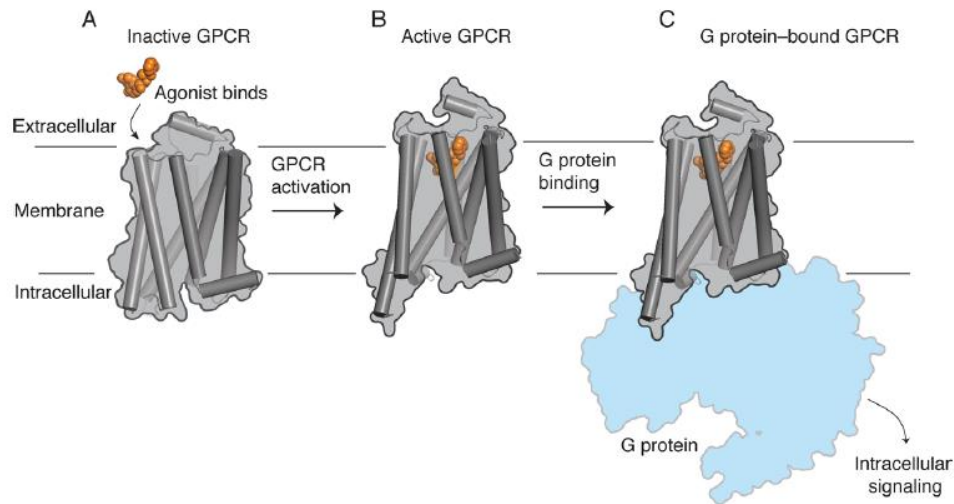


Figure 2: GPCR signaling: (A) an orthosteric ligand binds an inactive GPCR; (B) a ligand-bound GPCR undergoes a conformational change to its active state; (C) an active GPCR binds a G protein. Adapted from Latorraca, N. R., Venkatakrishnan, A. J. & Dror, R. O. GPCR dynamics: Structures in motion. *Chem. Rev.* 117, 139–155 (2017).

single signal would be wrong, since different signaling pathways can be modulated by a single GPCR and many GPCRs feature multiple sites for ligand binding that trigger transduction cascades in different ways.³ From a structural point of view, GPCRs are organized in a conservative fashion, featuring seven transmembrane α -helices connected by six loops, three of which protrude in the extracellular space and three in the intracellular space. From a dynamic point of view, GPCRs show a great variability of behaviors, mostly related to their different ability to change their conformational states equilibrium population upon ligand binding. In essence, modifications of the population of conformers of a GPCR occur in three ways:

- by altering the amount of time spent by a GPCR in a certain (active) conformation.
- by pushing the GPCR to adopt a conformation not present in the “normal” population of conformers.
- by increasing or decreasing the conformational switch rates.

These effects can be triggered not only by the binding of first messengers but, in a context of signaling modulation, also by intracellular partners (such as another GPCR or a G protein), by post translational modifications and by environmental conditions (pH). Conformational changes in GPCRs can affect also the extracellular loops and in these cases are coupled with transmembrane and intracellular conformational switches. Among these phenomena, of particular interest is the possibility to create binding sites for allosteric regulators that can modify the properties of GPCRs, including their affinity for primary messengers.

1.1.4 Transmission

Once the first messenger is received and converted in a biochemical signal, the cascade can proceed by means of what is considered “the heart” of the signaling machinery: the effectors. The effectors are all the enzymes that are involved in signal propagation, termination and modulation events, which are produced by:

- the interaction with and/or the modification of other effectors, thus regulating their activity and/or their subcellular localization
- the formation/degradation/release of second messengers

One of the most important features of the enzyme effectors is their possibility of being finely regulated by:

- conformational transitions (upon binding to effector molecules)
- post translational modifications that can
 - o directly regulate the activity of the enzyme
 - o trigger conformational modifications
 - o provide surface for the interaction with partner proteins
 - o regulate subcellular distribution
- membrane targeting

These phenomena provide flexibility to the enzymes that are involved in signaling pathways, allowing them to change their activity (no activity, low activity, high activity) and to serve as ideal tools for the transmission and the regulation of the signal.

1.2 The cAMP pathway

In 1971, Earl W. Sutherland won the Nobel Prize for his pioneering work that led to the discovery of one of the most complex and important signal transduction systems: the cAMP pathway⁴ (Figure 3). After nearly 50 years, this cellular system is still object of intensive research, in particular for what concerns its modulation by drug molecules as a strategy to treat a wide range of pathological conditions.

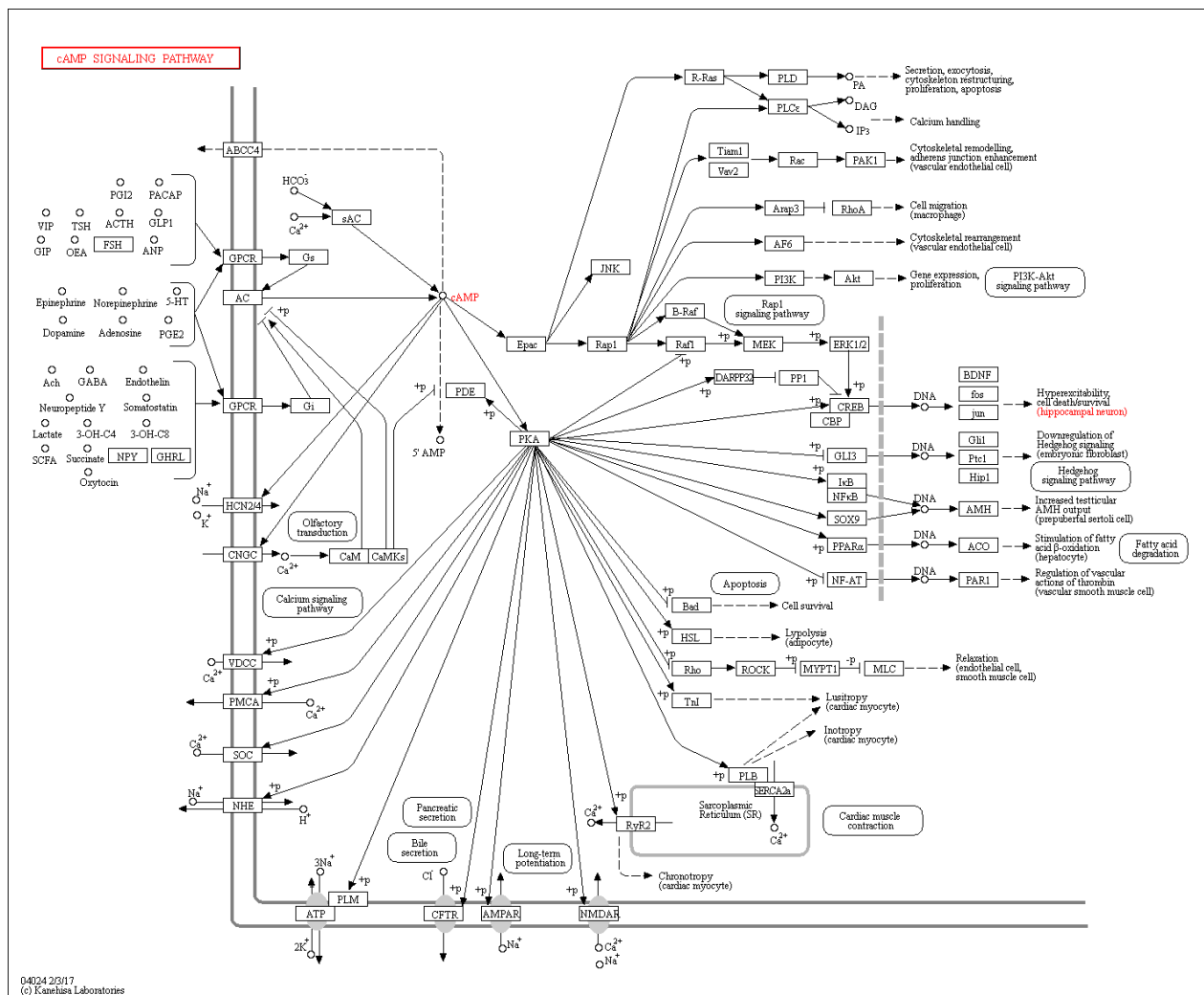


Figure 3: cAMP pathway - KEGG map04024

1.2.1 Synthesis

The intracellular levels of cAMP are regulated by a fine balance between its synthesis, which is catalyzed by adenylyl cyclases (ACs), and its removal, which is catalyzed by phosphodiesterases (PDE). Both of these enzyme families are characterized by a high variability of genes and isoforms, all of them displaying different expression patterns and a tight post-translational regulation.

The majority of adenylyl cyclases are associated to GPCRs. Upon binding of a ligand to the receptor, the α subunit of an $\alpha\beta\gamma$ G-protein complex, which by itself is able to bind and activate the AC, is released.

Starting from ATP, cAMP is produced through the deprotonation of the 3' hydroxyl group and a nucleophilic attack of the deprotonated oxygen towards the phosphorous.

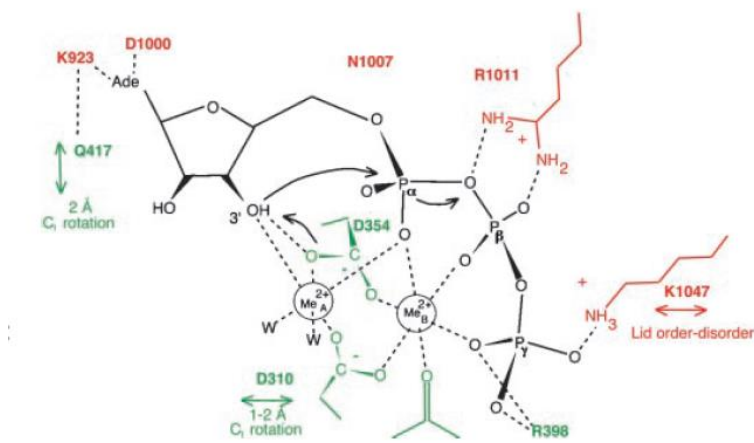


Figure 4 schematic representation of cAMP synthesis catalyzed by adenylyl cyclase

Key features of the catalysis are the stabilization of the transition state at the α -phosphate and the stabilization of the increased negative charge on the pyrophosphate leaving group, carried out by the two metal ions and by the residues Asn1007, Lys1047 and Arg1010⁵. The structure-function relationship of ACs have been extensively studied in

the last 10 years, thanks to up to 46 PDB entries. The structures of ACs in complex with ATP can be appreciated in PDB 4USW⁶.

1.2.2 Effectors

Upon reaching specific levels of cAMP within the cell, a number of effectors are activated. Among them, the most important are: Protein Kinase A (PKA), the guanine-nucleotide-exchange factor EPAC and the cyclic-nucleotide-gated channels.

PKA is a tetrameric complex composed by two catalytic subunits (C) and two regulatory subunits (R), both present with variable isoforms. This complex is activated upon binding of two cAMP molecules to both the R subunits. According to the traditional model of activation, the two R subunits dissociate from the complex generating a fully active enzyme. However, more recently, it has been shown that at a physiological cAMP concentration, PKA can function also as an activated but not dissociated holoenzyme⁷. The activity of PKA can be lowered by another partner protein: the protein-kinase inhibitor (PKI), which is also able to trigger the PKA C subunit export from the nucleus, thus decreasing its nuclear impact⁸. Like many other proteins involved in signal transduction, PKA can interact with anchoring partners (AKAPs) that specifically localize PKA near its targets or its modulators (close to ACs for fast activation and close to PDEs for signal termination) and that are responsible for PKA subcellular localization^{8,9}. Since the cAMP pathway is extremely complex and branched, it is not surprising that PKA has a huge number of cytosolic and nuclear targets, such as^{8,10}:

- metabolic enzymes like glycogen synthase and phosphorylase kinase, thus having a regulatory effect on glycogen synthesis and breakdown, or on acetyl CoA carboxylase for the inhibition of lipid synthesis.
- partners of other signaling pathways like phospholipase C or MAP kinases, thus in the first case inhibiting the partner and in the second case activating it.
- Raf and Rho, whose activity get reduced.
- ACs and PDEs, whose activity gets modulated
- transcription factors including cAMP-response element-binding protein (CREB), cAMP-responsive modulator (CREM) and ATF1. In these cases, phosphorylation allows the interaction of these proteins with their transcription coactivators (CREB binding protein and p300).

The Exchange-protein activated by cAMP (EPAC) exists in two isoforms (Epac1 and Epac2) codified by two genes. These effectors are characterized by a catalytic domain and by a regulatory one, which are able to bind RAS (RAS exchange motif and RAS association domain) and cAMP, respectively. As PKA, Epac proteins are involved in the regulation of a huge amount of cellular functions, ranging from cell adhesion to exocytosis, differentiation, proliferation and gene expression¹¹.

One last effector within the cAMP signaling pathway is the family of cyclic-nucleotide-gated ion channels (CNG), which transports cations with poor selectivity upon binding to four cAMP (or cGMP) molecules. These gates are particularly important in photoreceptor and in sensory neurons¹².

1.3 cAMP-Specific Phosphodiesterase

Negative modulation of the cAMP signal occurs through the decreasing of cAMP levels and through the counterbalancing of cAMP positive effectors. cAMP levels are modulated by its removal by cAMP-specific phosphodiesterases (primary PDE4s) and by the decreasing of its synthesis via ACs deactivation. At the same time, positive effectors like PKA are counterbalanced by the activity of other actors like phosphatase PP1 and PP2A⁸.

The signal termination driven by PDE4s occurs through the hydrolysis of cAMP. This process yields the non-cyclic version of the molecule, which is no longer capable of delivering a signal.

1.3.1 Coordination and catalysis

The catalytic site of PDE4s is characterized by two metal ions: Zn^{2+} and Mg^{2+} . The first metal is coordinated by two histidines and two aspartates (H330, H366, D367 and D484) while the second one is coordinated by one bridging aspartate (D367). Both coordination geometries are octahedral: all of the free coordination positions are saturated by water molecules and a hydroxyl group bridges the two metals in the cAMP-free enzyme (Figure 5). The structure of the PDE4B-AMP complex¹³ reveals that a phosphoryl oxygen atom of AMP displaces the hydroxyl ion and becomes the interacting partner for both the metal ions, thus clearly suggesting that the OH^- acts as the nucleophile in the hydrolysis of the phosphodiester bond¹⁴. The subsequent protonation of O3 by His326 generates the leaving group. The importance of this residue is highlighted by its conservation among different PDE families; the same holds for its interacting partners, Tyr325, Glu505 and His370, which are also totally conserved.

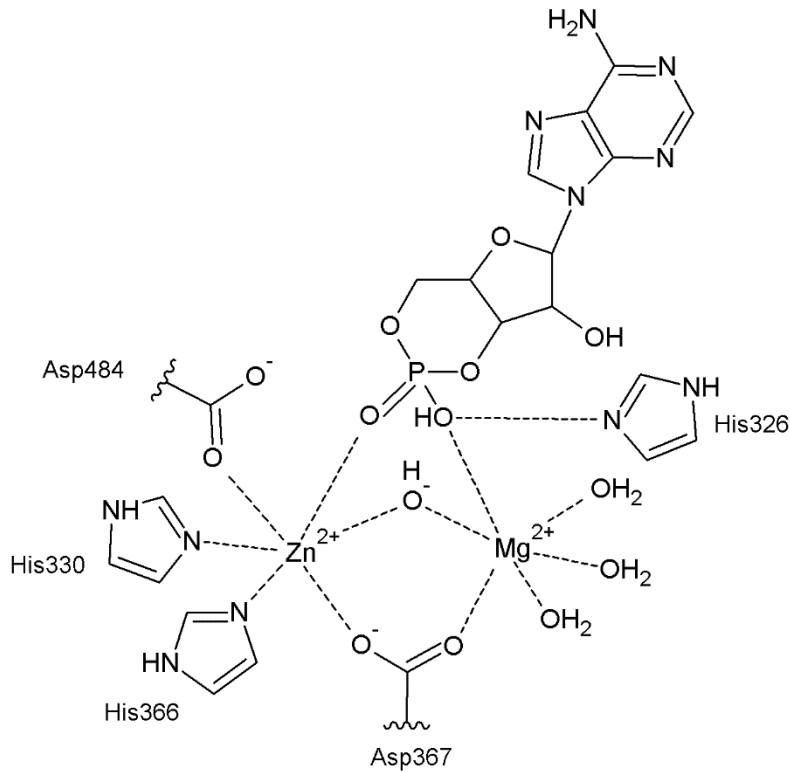


Figure 5: Metal ion coordination scheme in the PDE4-cAMP complex

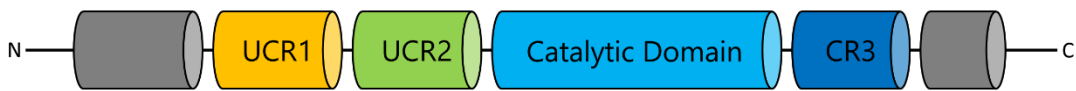
1.3.2 PDE4 gene structure

In mammals, PDE4s are encoded by four genes, each of them being processed by alternative splicing and yielding more than 20 isoforms. Moreover, isoform-specific promoters within each gene are responsible for their distinct expression patterns in different tissues and within the central nervous system^{15,16}.

- PDE4A
 - Location: 19p13.2
 - Isoforms: PDE4A 1-7
- PDE4B
 - Location 1p31.3
 - Isoforms: PDE4B 1-5
- PDE4C

- Location: 19p13.11
- Isoforms: PDE4C 1-3
- PDE4D
 - Location: 5q11.2 – q12.1
 - Isoforms: PDE4D 1-12 (no 4 and 10)

The primary difference between the different isoforms that are generated by alternative splicing is the length of the transcripts. As PDE4s are modular enzymes, this results in the presence or in the absence of functional modules that are able to confer specific properties to the isoforms.



The different modules of this system are:

- Catalytic domain: independent module that retains the catalytic function of the enzyme. It is highly conserved among different PDE4s
- UCRs: these modules, which are located at the N-terminus of the catalytic domain, are involved in the regulation and dimerization of the enzyme. Despite being relatively conserved among PDE4s, they feature some crucial differences that are particularly interesting for drug-discovery purposes.
- CR3: this domain is composed by a single and relatively short α -helix (10-aminoacid circa) that is also involved in the regulation of the enzyme.
- terminal regions: these are isoform-specific sequences that are responsible for the interaction with other protein substrates

PDE4 isoforms can be divided in super-short, short and long forms, depending on their length. The super-short isoform displays only the catalytic domain and a truncated version of UCR2, short isoforms carry a complete UCR2 while long isoforms display also UCR1 and the Terminal unique regions.

1.3.3 Dimerization

The modularity of PDE4s is reflected in their different expression, activity, regulation and structural properties. One of the main differences among isoforms, which influences their other

cited properties, is their ability to form homodimers. By and large, long isoforms (such as PDE4D3) are functionally dimeric while short (such as PDE4D2) and super-short isoforms are monomeric. This categorization was introduced and validated by several studies using co-immunoprecipitation, gel-filtration, density-gradient centrifugation and yeast-two hybrids assays carried on the wild-type isoforms and on mutants¹⁷⁻¹⁹.

From a structural point of view, three contact areas are involved in dimerization: the first is between two catalytic domains, the second between UCR1-2 modules and the third between the UCR2 domain of one monomer and the catalytic domain of the other.

The first interaction surface is the smallest in the dimer (buried area of 700 Å²); it is shared by all PDE4s, as can be appreciated in all the available crystal structures in the PDB, albeit not sufficient to form a dimer in solution¹⁶.

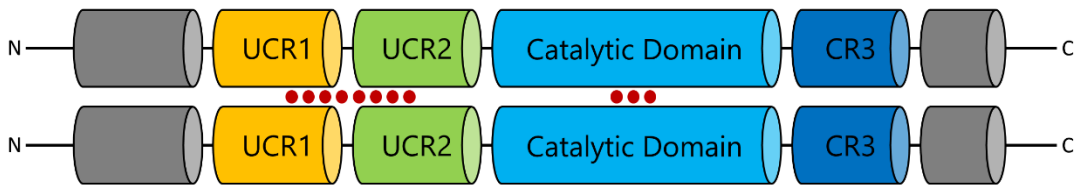


Figure 6: PDE4 protein scheme and dimer interaction surfaces

The two catalytic domains interact by means of a loop (H-loop) with a central hydrophobic site (Tyr389 and Met388), multiple hydrogen bonding interactions (Asn390, Asn397) and salt bridges (Asp391-Arg427)²⁰.

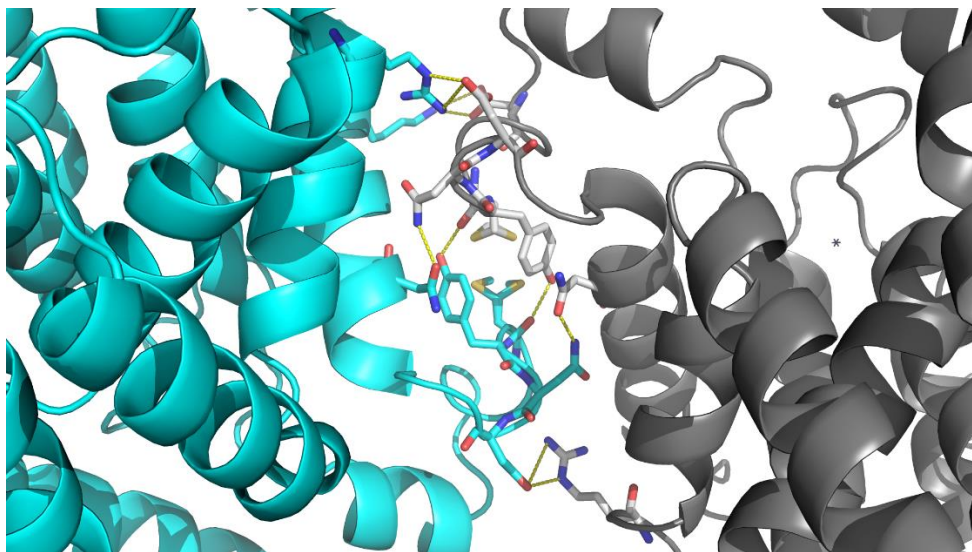


Figure 7: H-loops interacting in all the PDE4D catalytic domain crystal structures

The second interaction surface is composed by a wide portion of the PDE4 regulatory domain, which includes the C-terminal part of UCR1 and the N-terminal part of UCR2. This second dimerization domain completes the full set of interactions that are necessary for the enzyme to form a dimer and has been recently structurally characterized on a PDE4B long isoform (PDB code: 4WZI)²¹. It is shaped as a four-helix antiparallel bundle, featuring a core of conserved hydrophobic residues that generate a total buried area of 1300 Å².

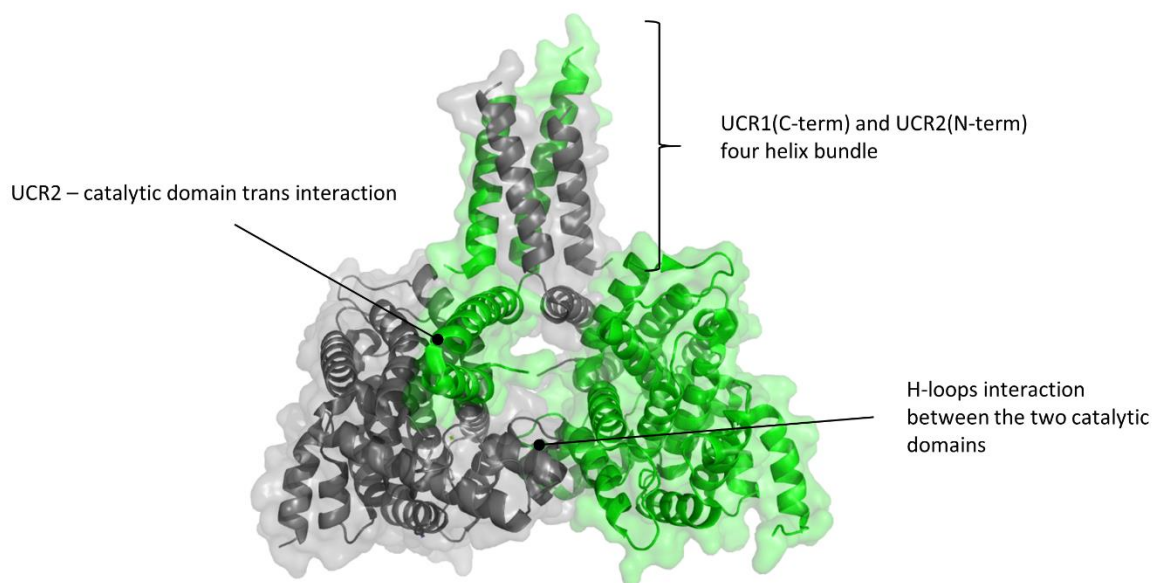


Figure 8: PDE4B cross-linked dimer. PDB ID: 4WZI

The last interface between two monomers is formed at the level of the second and third helices of UCR2, in a cross-interaction with the catalytic domain of the adjacent monomer, creating a buried area of 950 Å². UCR2 acts as a lid of the catalytic pocket occluding its entrance, thus generating an auto-inhibited conformation that has gained attention for drug design purposes and that will be discussed in the next chapters.

1.3.4 HARBS and LARBS

Aside from their structural features, the differences between monomeric and dimeric PDE4s are reflected in their enzymatic and pharmacological properties^{16,18}. The most updated models describe the dimeric isoforms featuring UCR2s domains that close the entrance to the catalytic pocket reversibly, limiting either access of the substrate or product release. Soaking experiments and kinetic data on a “fully closed” stabilized enzyme demonstrated that when UCR2 adopts the closed conformation, the active site is still accessible by substrate and inhibitors. The K_m of this construct is similar to that of the activated enzymes, while its specific activity is reported to be 10 to 60 times lower²¹.

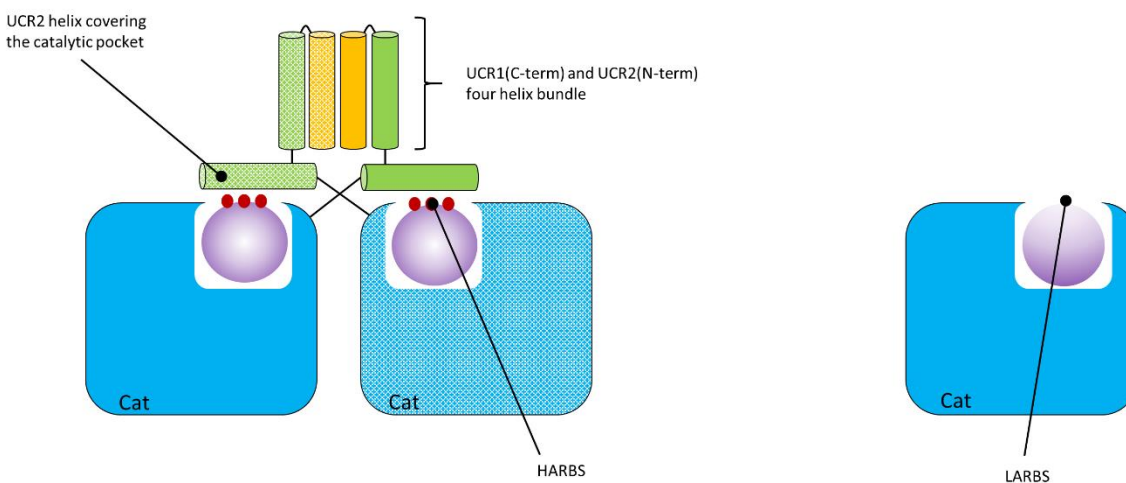


Figure 9: Schematic representation of HARBS in long isoforms (left) and LARBS in short isoforms (right)

From a pharmacological point of view, the existence of a closed conformation in dimeric isoforms generates what is generally referred to as HARBS (High Affinity Rolipram Binding Site) and LARBS (Low Affinity Rolipram Binding Site)^{16,18,22–25}. In principle, some PDE4 inhibitors, such

as the standard Rolipram, display two different affinities for the targets, one characterized by a lower IC_{50} (HARBS) and one by a higher IC_{50} (LARBS). For instance, Rolipram has been reported to inhibit PDE4D3 (dimeric) and PDE4D2 (monomeric) with an IC_{50} of 1.17 and 0.55 μ M, respectively²⁶. It is important to note that these values are strongly dependent on the assay being used, on the specific protein construct, on the phosphorylation pattern of the protein and on the experimental conditions in which the assay is done. Therefore, as different values are reported in the literature, a correct comparison is only possible when considering the same experimental conditions. In the case of long PDE4 isoforms, which are able to form dimers, both binding states are possible. This is due to the presence of UCR2 in a closed conformation, which can assist the binding of the inhibitor. Conversely, monomeric isoforms show only LARBS states, in which the inhibitor binds only to the catalytic domain of PDE4.

1.3.5 PDE4D Regulation by phosphorylation

PDE4s are regulated by a complex pattern of isoform-specific phosphorylation states, involving the activity of a broad range of phosphorylating enzymes^{27,28} as well as numerous signaling pathways. In the following, the focus will be on the regulation of PDE4D, which is the object of this work.

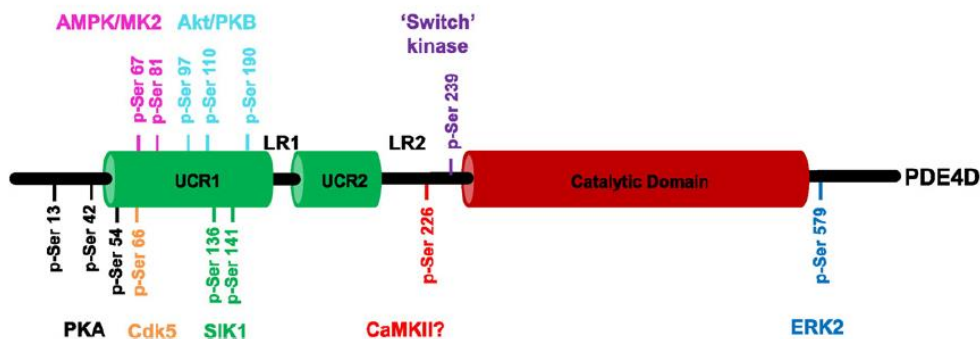


Figure 10: Schematic diagram of PDE4 phosphorylation sites. From: Mika, D. & Conti, M. PDE4D phosphorylation: A coincidence detector integrating multiple signaling pathways. *Cell. Signal.* 28, 719–724 (2016)

1.3.5.1 PKA phosphorylation

PDE4 long isoforms share a RRESF motif located in the N-terminal portion of UCR1. This motif hosts Ser54 (numbering referred to PDE4D), which is one of the most important phosphorylation sites for PDE4 regulation. This post-translational modification is introduced by PKA and yields an increase in enzyme activity (V_{max}) and an increase in affinity for Mg^{2+} ^{28–30}. Although there is a

general consensus that phosphorylation provides its functional effects through a mechanism that involves a conformational rearrangement of the UCR motifs, a precise mechanistic picture of this phenomenon is still matter of investigations²⁸. It has been hypothesized that phosphorylation may affect the orientation of UCR1, which could fold back on UCR2, thus modifying its position over the catalytic domains^{21,31}. Since phosphorylation of Ser54 has also been reported to increase the affinity for Rolipram³⁰ (towards a possible HARBS configuration), it is probable that UCR2 is not totally removed from the capping position, resulting in a model that includes a still active closed conformation (of a long isoform), albeit with a significantly decreased Vmax²⁸.

Additional phosphorylation sites for PKA are Ser13 of PDE4D3 and Ser42 of PDE4D7. The first one is reported to be an activating phosphorylation that increases affinity of PDE4D3 for muscle AKAPs³², while the second is a deactivating phosphorylation that occurs under basal conditions³³.

The regulation of PDE4 by PKA takes place in localized cellular compartments where AKAPs (A Kinase Anchoring Proteins) coordinate the complex for PKA activity on PDE4s, generating the so-called “signaling-microdomains”³⁴.

1.3.5.2 ERK2-MAPK phosphorylation

Phosphorylation by the Extracellular-signal Related Kinases (ERK, p42MAP) is quite common among PDE4s. ERK2 phosphorylates PDE4D3 on Ser579 in the C-terminal portion of its catalytic domain, causing its inhibition³⁵. Based on the structural characterization of the UCR domains (described in chapter 1.3.3), the effect of this mutation has been hypothesized to be delivered upon a stabilization of the UCR2 capping helix over the catalytic domain. In fact, this phosphorylation has been shown to inhibit only long isoforms^{21,31}.

In conditions of high oxidative stress, the Ser579 phosphorylation on PDE4D3 can be reprogrammed by a second phosphorylation (by PI3 kinase) on Ser239. This phenomenon reverts the first modification, causing an increase of hydrolytic activity that will promote an inflammatory response³⁶. Since Ser239 is found in a disordered portion between UCR2 and the catalytic domain, its mechanistic involvement in the activation of the enzyme is still elusive²⁸. A similar reversion mechanism has been described for PDE4D9 activation during G2/M mitotic transition, where an unknown switch kinase is able to revert Ser585 (Ser579 in D3 isoform) inactivation²⁸. Other more elusive phosphorylation sites have been described for PDE4D9, such as for instance Ser67, which attenuates enzyme activation by PKA.

1.3.5.3 CaMKII

It has been shown that in the context of excitation-contraction coupling in cardiac myocytes, CaMKII takes part in the cAMP regulation under basal conditions and also upon β -stimulation, by regulating PDE4D3 activity. CaMKII has been shown to phosphorylate PDE4D3 *in vitro*, to activate endogenous PDE4s *in vivo*, and to enhance PDE4D3 activation by PKA³⁷. While the exact phosphorylation site for CaMKII has not yet been found experimentally, the most probable site has been recognized computationally in LR2²⁸.

1.3.5.4 Other phosphorylation partners

The complex network of pathways in which PDE4s are involved and regulated is still matter of investigation. Indeed, other phosphorylating partners have been discovered and their characterization is currently ongoing²⁸:

- Cdk5 has been recently demonstrated to phosphorylate some PDE4s (including PDE4D) in UCR1. This phosphorylation is thought to act as a basal modification that works as a priming factor for PKA³⁸.
- Protein kinase B (PKB) has been reported to phosphorylate PDE4D in three possible sites (Ser97, Ser110 and Ser190) in bronchial smooth muscles. It has been shown that arctigenin is able to reduce PKB phosphorylation, participating in a synergistic relaxing effect in bronchorelaxation³⁹.
- PKA activity, stimulated by cAMP, increases insulin secretion by pancreatic β -cells. In this context, an increase of cAMP levels stimulates salt-induced kinase 1 (SIK1) expression. PDE4D has been shown to be phosphorylated and activated by SIK1 on Ser136 and Ser141, decreasing cAMP levels and therefore insulin secretion⁴⁰.

1.3.6 PDE4D localization and interactions

In early cellular signaling studies, one of the biggest challenges was to understand how ubiquitous second messengers such as cAMP could trigger highly specific responses, despite activating a huge variety of receptors and having their levels regulated by a number of partners. Now, we know that the levels of second messengers are finely tuned not only across the whole cell, but also in microdomains. In the case of cAMP, microdomains are established by ACs and PDEs that are tethered to subcellular compartments by means of direct interactions with membranes, or through

other scaffolding proteins⁴¹. The list of PDE4s interactions that have been identified to date is by no means exhaustive; here, only the best characterized mechanisms will be presented.

As previously mentioned, AKAPs represent one of the major scaffolding protein for cAMP microdomains management, featuring a conserved domain for the interaction with PKA and domains for the interaction with other proteins such as PDE⁴¹. The best known examples of AKAPs involvement in cAMP microdomains regulation are⁴¹:

- mAKAP organizes PDE4D3 and PKA in the perinucleus, regulating cardiac myocytes. High levels of cAMP activate PKA bound to mAKAP, triggering the activation of PDE³⁴.
- AKAP40 binds PDE4D3 in centrosomes. This complex works in the same way as in the perinucleus⁴².
- AKAP18 δ interacts with PDE4D3 and PKA, tethering them to AQP2-bearing vesicles. In this pathway, antidiuretic hormone triggers the distribution of water from vesicles to the plasma membrane of some renal cells. The complex maintains a low basal level of cAMP, preventing water reabsorption⁴³.

PDE4s interact with plasma membrane receptors. Examples are provided by PDE4A4 and PDE4A5, which interact with p75 neurotrophin receptor (p75NTR) in a mechanism that regulates the equilibrium between tissue repair and scar formation^{41,44}.

Localization of PDE4 can also be a function of the cell state. In HEK293, the activation of β -adrenoreceptors triggers an increase in cAMP levels, which are subsequently lowered by a desensitizing binding of β -arrestin to the receptors. In a resting state condition, β -arrestin binds PDE4D5 in the cytosol and when it is tethered toward the receptors, it carries PDE4D5 along, which will lower local cAMP levels, contributing to signal termination^{41,45}.

Another important scaffolding protein for PDE4 is RACK1 (Receptor of Activated C Kinase 1). RACK1 acts as a hub for protein complex formation. It can be found in the cytosol, ribosomes and plasma membranes and it is involved in a broad range of mechanisms: virus infection, cell migration, neuronal development and angiogenesis⁴¹. PDE4D5 is able to bind both β -arrestin and RACK1, but in a mutually exclusive fashion. In HEK293, it has been demonstrated that there is an equilibrium between the two interactions and that this equilibrium can be impaired by knocking

down RACK1. In this case, the population of PDE4D5 bound to β -arrestin increases, and too much phosphodiesterase activity will be provided within the β -adrenoreceptors environment^{41,46}.

As previously mentioned, the characterization of all the molecular interactions involving members of the PDE4 family is not exhaustive and several other partners and mechanisms are currently being investigated, such as, for instance, the interactions with myomegalin, shankr, Disc1, Src proteins and many others⁴¹.

1.3.7 The pocket

The PDE4D active site is divided in three distinct areas (known as the M-pocket, the S-pocket and the Q-pocket) that, together, generate a 15Å-deep cavity, with an entrance of 20x15Å and a total volume of 925 Å³ ⁴⁷.

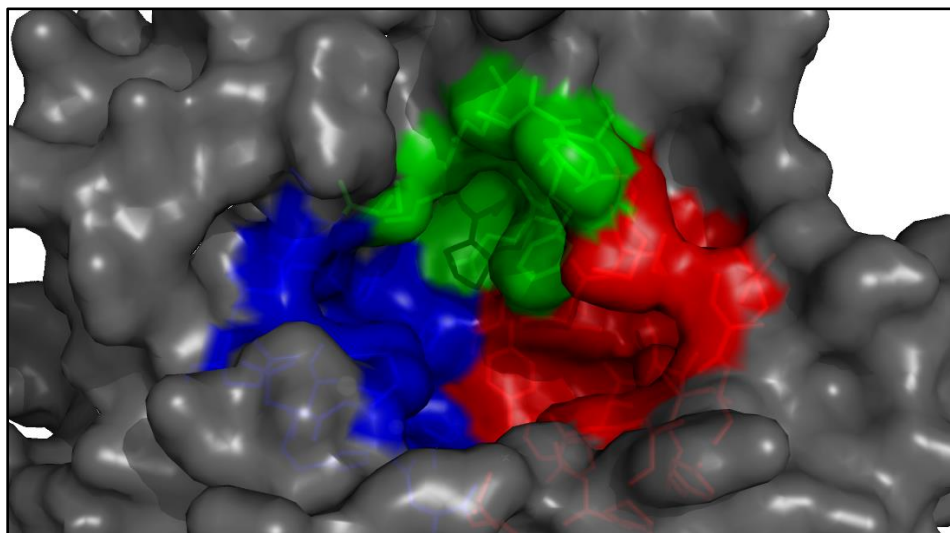


Figure 11: PDE4D catalytic domain, view from above the entrance. M-pocket (Blue), Q-pocket (Red), S-pocket (Green). PDB ID: 6F8T, adapted from Prosdocimi et al. 2018. Biochemistry 57 2876-2888

<i>Name</i>	<i>Notes</i>	<i>Residues</i>
M-pocket	Contains the dimetal cluster (Zn^{2+} and Mg^{2+}) and is the most conserved among PDEs	H396, H330, H366, D367, H370, N375, L395, E396, D438, M439, D484
S-pocket	Mainly featuring polar residues, in most of the inhibitor complexes is filled with water molecules.	G372, S374, E505, F506, Q508, S521, C524

Q-pocket	Most important domain for small molecule recognition. It can be further divided in 3 micro-regions: 1) a saddle, defined by a glutamine residue that is extremely important for hydrogen bonding with ligands, 2) a P-clamp region, which creates a hydrophobic groove, 3) two hydrophobic micro-pockets (Q ₁ and Q ₂), which are separated by the glutamine saddle	G372, S374, E505, F506, Q508, S521, C524
----------	--	--

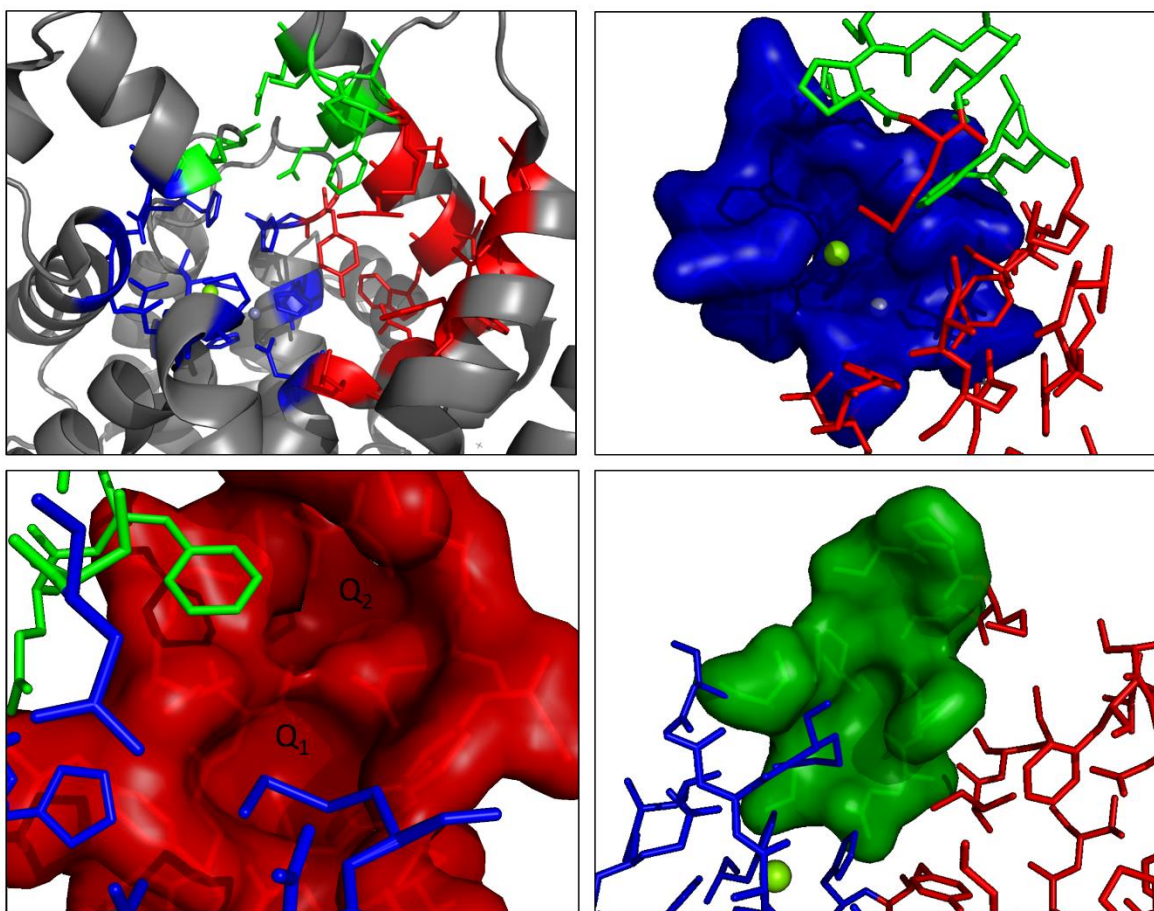


Figure 12: Details of the active site of the PDE4D catalytic domain. Ball and stick representation of the full pocket (top-left), M-pocket (top-right), Q-pocket (bottom-left), S-pocket (bottom-right). PDB ID: 6F8T, adapted from Prosdocimi et al. 2018. *Biochemistry*. 57 2876-2888

As pointed out in the previous chapters, long PDE4 isoforms display a complex regulatory machinery in which a portion (UCR2 domain) is able to specifically cover the entrance to the

catalytic domain. Since this phenomenon is intrinsically transient and involves flexible domains, a perfect structural characterization of this interaction at the atomic level is still elusive. To date, in the PDB the structures in which UCR2 is shown to cover the catalytic site are either only partial^{31,48} or have been obtained using mutagenic constructs that stabilize a specific conformation²¹. Moreover, UCR2 is not the only domain that has been shown to be able to act as a lid for the pocket. In fact, CR3, a helix at the C-terminus of the catalytic domain can also work in the same way⁴⁸.

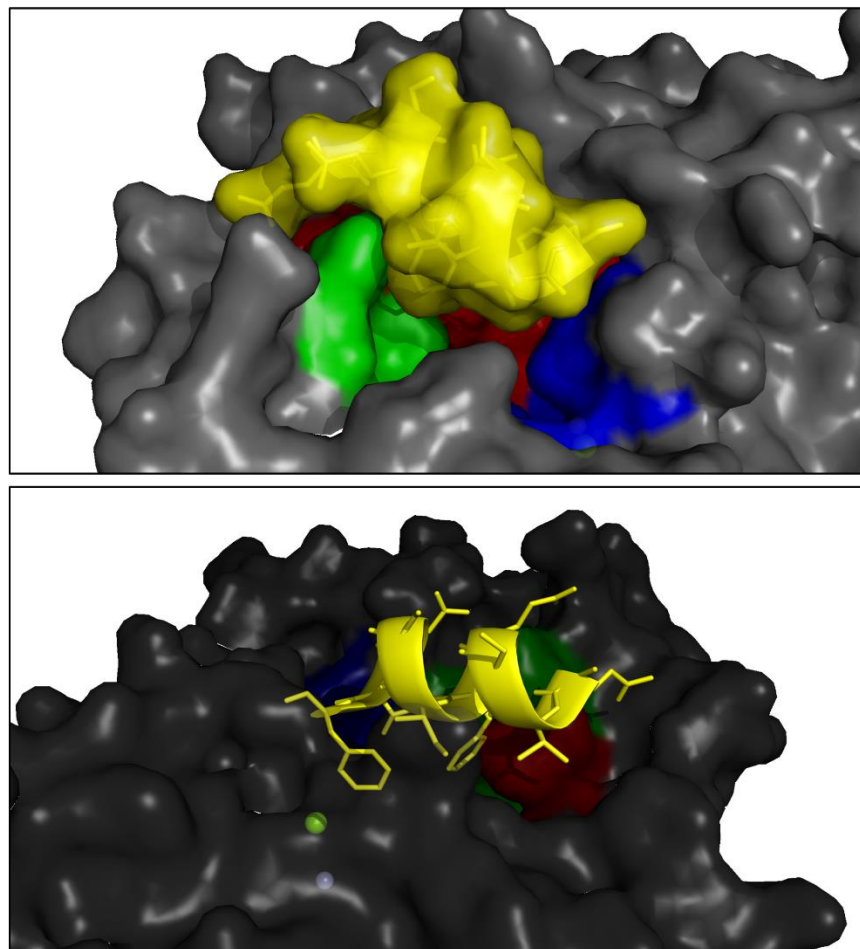


Figure 13: UCR2 (Yellow) positioning on PDE4D catalytic domain. PDB ID: 3IAD

The UCR2 helix is stabilized on top of the catalytic domain by means of both polar and hydrophobic interactions. Gln192 (UCR2) is placed at the border of the pocket where it can participate in a hydrogen bond with the Asn528 backbone (catalytic domain). Phe196 and Phe201 (UCR2) dive into the hydrophobic groove defined by Ile542, Met439 and Leu485, while Val193, Ile197, Leu202 (UCR2) interact with Gly537, Tyr541, Ile542 and Met443 (catalytic domain)³¹.

The presence of a capping helix on the entrance of the active site generates an expanded pocket that has gained the attention of the scientific community. In particular, the pioneering work done by Burgin and colleagues provided evidence that it is possible to exploit the hydrophobic residues Phe196 and Phe201 of the lid to develop allosteric and specific modulators for the PDE4D/B long isoforms. This strategy has led to the development of some of the most promising PDE inhibitors for memory and cognitive-related diseases, such as Alzheimer's and Fragile x-syndrom^{31,48-51}.

1.4 Phosphodiesterase 4 as pharmaceutical target

Given the extensive involvement of PDEs in the regulation of such a large spectrum of cellular processes, it is not surprising that these enzymes are considered important pharmaceutical targets. While the development of PDE inhibitors has always been intrinsically complex due to the side effects associated to this kind of therapy, efforts in this direction have intensified when the possibility of selectively targeting different PDE isoforms became more evident^{50,52-55}.

Most of the information on the effects of PDE4s inhibition has been acquired by comparing PDE4(x) knockout mice with PDE(x) inhibited mice (Table 1).

Table 1: Phenotypes associated with PDE4 knockout mice. Adapted from Miles D. Houslay. Phosphodiesterase-4 as a therapeutic target. 2005. Drug Discovery Today. Abbreviations: NR, not reported. /, normal level

<i>Phenotype</i>	<i>Response for PDE4A knockout</i>	<i>Response for PDE4B knockout</i>	<i>Response for PDE4D knockout</i>
Neonatal growth	/	/	impaired
TNF- α production by leukocytes stimulated by LPS	/	90% decreased	/
Airway hyperactivity in response to allergens and cholinergic agonist	Reduced	Reduced	/
IL-4, IL-5, in bronchoalveolar lavage fluid mediated by allergens	NR	Reduced	/
Eosinophil recruitment to bronchoalveolar lavage fluid mediated by allergens	NR	64%	/
Neutrophil recruitment to bronchoalveolar lavage fluid mediated by LPS	/	31% decreased	48% decreased

Tracheal contractility induced by cholinergic agonists	/	/	34% decreased
Shortening of α 2 adrenoreceptor anesthesia (correlates to emesis)	NR	/	reduced
Cognition	NR	/	Increased
Antidepressant	NR	Increased	Increased
Anxiolytic	NR	Decreased	/

1.4.1 Side effects

The major problems encountered by PDE4 inhibitors in reaching the market are not related to low desirable effects, low response in humans, adsorption or pharmacokinetics difficulties, but to the existence of considerable side effects (mainly emesis and diarrhea). The increased neuronal activity within the area postrema in the CNS is associated with nausea and emesis, which is one of the most severe drawbacks of PDE4 inhibition. Indeed, this occurs both when 8-bromo cyclic AMP is systemically administrated or when PDE4s are strongly inhibited^{56,57}.

The connection between PDE4 inhibition and emesis was established 20 years ago. Subsequently, for many years scientists have struggled (and this is still partially a matter of debate) to understand which isoform causes such problem, also because proteomic studies revealed that both PDE4D and PDE4B are present in the area postrema^{57,58}. The majority of isoform-specific effects were studied in rodents, which unfortunately lack the emetic reflexes. In this animal model, it is possible to exploit a biological effect that correlates with emesis: the reversal of xylazine-induced anesthesia triggered by PDE4 inhibition⁵⁹. Today, we know that, most likely, the main actor in the emesis response is the PDE4D isoform^{59,60}. Indeed:

- deletion of PDE4D (and not of PDE4B) reduced the xylazine-induced hypnotic effects.
- the ability of PDE4 inhibitors to shorten anesthesia is impaired when PDE4D is knocked down, while it is still measurable when PDE4B is knocked down

These studies are still considered not completely conclusive because of the concomitant presence of several other studies that report controversial results⁵⁷. However, generally speaking, PDE4B has become the preferred target for airway diseases (such as for instance COPD and asthma), while PDE4D remains in competition with PDE4B for what concerns the treatment of memory and

cognition related diseases (such as for instance Alzheimer's disease), where the side effects associated with its inhibition are counterbalanced by a better performance in cognition amelioration with respect to PDE4B inhibition^{50,51,54,55,61}.

1.4.2 Airway disease

In Chronic Obstructive Pulmonary Disease (COPD), inflammation is triggered by macrophage activation, chemokine secretion, neutrophils and CD8+ lymphocytes recruitment. Moreover, macrophages and lymphocytes secrete proteases that worsen airway fibrosis, increase mucus secretion, and ruin the alveolar wall^{57,62}. In mice, it is possible to partially model COPD by administering endotoxin, which recruits neutrophils. This phenomenon is inhibited by 50% in both PDE4B and PDE4D knockout mice and the same result can be reproduced in PDE4B/D inhibited wild type mice⁶³. Until 2017, those PDE4 inhibitors that were developed for respiratory pathologies were the ones that reached phase III clinical trials and the market. Many clinical trials certified the potential of this therapy, in particular for COPD^{57,64}. Research and development related to this kind of PDE inhibitors is still ongoing.

1.4.3 Central Nervous System

The role of PDE4s inhibition in the nervous system aims at the potentiation of the cAMP cascade that triggers the activation of those transcription factors, CREB in particular, that enhance the expression of genes connected to learning and memory^{53,54,61,65,66}. PDE4s differential expression reflects the different isoform-specific roles; among all the isoforms, PDE4D has been shown to be the most expressed in hippocampal CA1⁵⁸. Knockout mice for PDE4D and PDE4 inhibited mice have shown improvement of long-term hippocampal memory. The effects of Rolipram, which to date is considered the gold-standard PDE4 inhibitor, ranged from the amelioration of memory related tasks (such as objects recognition) to the passive avoidance learning and prefrontal cortex-dependent memory^{61,67,68}. Beneficial effects of Rolipram were also reported in animals with induced memory deficits^{68,69}.

PDE4B and PDE4D are two highly similar PDE4 isoforms, both of them involved in cAMP modulation in the CNS. As mentioned before, PDE4D is believed to be the most promising target due to the memory-enhancing effects of its inhibition, while PDE4B has been proposed as a target with less side effects. However, it is worth noting that PDE4B inhibition is not the only strategy pursued for reducing side effects. Other advances in this regard will be described in the next

chapter. Recently, a comparison between the pharmacological profiles of the selective PDE4D/B inhibition was published by Zhang and colleagues⁵⁰. Interestingly, they reported that while both strategies produce increased CREB phosphorylation in HT-22 cells, they behave differently when tested *in vivo*. Indeed, PDE4B inhibition has shown antidepressant-like effects in mouse forced swim tests (FST) and tail suspension tests (TST) while PDE4D inhibition has shown improved cognition in mouse Norwegian tenecteplase stroke (NOR) tests. These findings suggest that both isoform specific inhibitions can be exploited separately for different purposes.

The importance of PDE4D in cognition is also highlighted by a rare genetic disease, Acrodysostosis without hormone resistance (ACRDY2). This genetic disorder, which causes mental retardation, brachydactyly and facial dysplasia, is caused by several missense mutations of PDE4D. Interestingly, the mutations do not affect the catalytic ability of the enzyme, but they are located 1) at the interface between the catalytic domain and the regulatory domains and 2) in the PKA phosphorylation site Ser129, most probably impairing the regulation mechanism of the enzyme^{49,70,71}.

1.4.4 Alzheimer's disease

The promising effects of PDE4 inhibition in CNS are counterbalanced by severe side effects. For this reason, PDE4 inhibitors have been historically proposed as a possible treatment for highly problematic pathological conditions (such as for instance Alzheimer's disease) for which a solution is not yet available. In fact, the lack of an efficacious treatment for Alzheimer's disease (which would mainly involve targeting A β -amyloid plaques) stimulated scientists to develop new non-amyloid-based strategies. After the early work by Zhang and colleagues⁵⁰ (and other more recent publications), exposed in the previous chapter, PDE4D is now definitely considered as a key pharmaceutical target for cognition improvements.

The preclinical development of new and safe PDE4 inhibitors for Alzheimer's has gained attention from the scientific community. The two main strategies that are being investigated are the allosteric modulation³¹ and the development of inhibitors that show specificity for the PDE4D sub-isoforms⁷². However, some compounds that also belong to other inhibitor classes (such as PDE4 pan-inhibitors) have already reached the clinical phases⁷³:

Drug	Company	Notes	Studies
MEM 1414	Roche	Roche stopped the clinical phases and acquired the previous owner of the patent. Details unknown	
MK-0952	Merck		Clinical I completed but no results disclosed. Identifier: NCT00362024
HT-0712	Dart Neuroscience	Further studies were carried out but the results remained undisclosed.	Clinical II completed: improvements in long-term memory in patients with memory impairments connected to aging. Identifier: NCT02013310
BPN 14770	Astrazeneca		Clinical I (first) completed: not effective. Identifier: NCT02051335. Clinical II completed: Improvements in cognition of healthy adults. Identifier; NCT01433666.
BPN 14770	Tetra Discovery Partners	Negative allosteric modulator	Clinical I (two studies) and further investigations: safe, well tolerated but with headache as main adverse event. Improvements in working memory. Identifiers: NCT02648672, NCT02840279

1.4.5 X-fragile

Fragile-X Syndrome (SFX) is a genetic disorder that causes intellectual disability and problems belonging to the autism spectrum. The molecular cause of SFX is a mutational inactivation of the gene coding for FMRP (Fragile X Mental Retardation Protein), which is a protein associated with the regulation of mRNA translation in pre- and post-synapsis. The absence of FMRP results in the

overexpression of proteins whose levels should, instead, remain perfectly balanced, such as ion channels, signaling and structural proteins^{74,75}. The possibility to target the X-fragile disorder by inhibiting PDEs was initially contemplated at the time when Berry-Kravis and colleagues reported reduced cAMP levels in FXS-affected cells⁷⁶. Since then, *Drosophila* has been used as the animal model to study PDE4 inhibition in SFX. However, *Drosophila* features only a single PDE4 gene that does not possess the UCRs domains that distinguish PDE4 from other PDEs and also all the different PDE4 isoforms. In the last couple of years, PDE4D specific allosteric modulators (see next chapter) have been tested in mice models in order to assess the therapeutic benefits of a sub-isoform specific targeting strategy⁵¹. A cutting-edge compound developed by Tetra Discovery Partners (BPN14770) is now in phase II clinical trials for X-fragile syndrome (Identifier: NCT03569631).

1.5 Phosphodiesterase 4 inhibitors

1.5.1 Structural bases

The structural bases that define the nature of the interactions between small molecules and phosphodiesterases (with particular attention to PDE4s and PDE5s) have been studied and extensively reviewed by Zhang and colleagues in 2004⁴⁷.

Most of the inhibitors have in common an anchoring core that is characterized by two main interactions:

- an aromatic ring sandwiched in a P-clamp. The site is composed by two hydrophobic residues that act as two jaws, creating a docking groove. While the P-interacting residue is a conserved phenylalanine (Phe538 in PDE4D3), the other one varies between the different PDE family members (Ile502 in PDE4D3)
- a hydrogen bond with a conserved glutamine residue (Q535 in PDE4D3). This residue (also called Q-switch) is able to change its orientation upon cAMP recognition and the sub-pocket in which it is localized is named after it.

The binding is then completed by other hydrophobic interactions, water-mediated hydrogen bonds and direct hydrogen bonds to protein residues.

In the history of PDE4 drug development, the biggest family of compounds feature a catechol moiety as the anchoring core with two substituents on the catecholic oxygens and an R group on the other side of the phenyl ring (see Figure 14). The two substituents on the catecholic oxygens are variable hydrophobic groups, such as cyclopentyl, cyclopropyl or propyl on one side, and generally a methyl on the other side. These groups occupy the Q₂ and Q₁ sub-pockets, respectively (see paragraph 1.3.7).

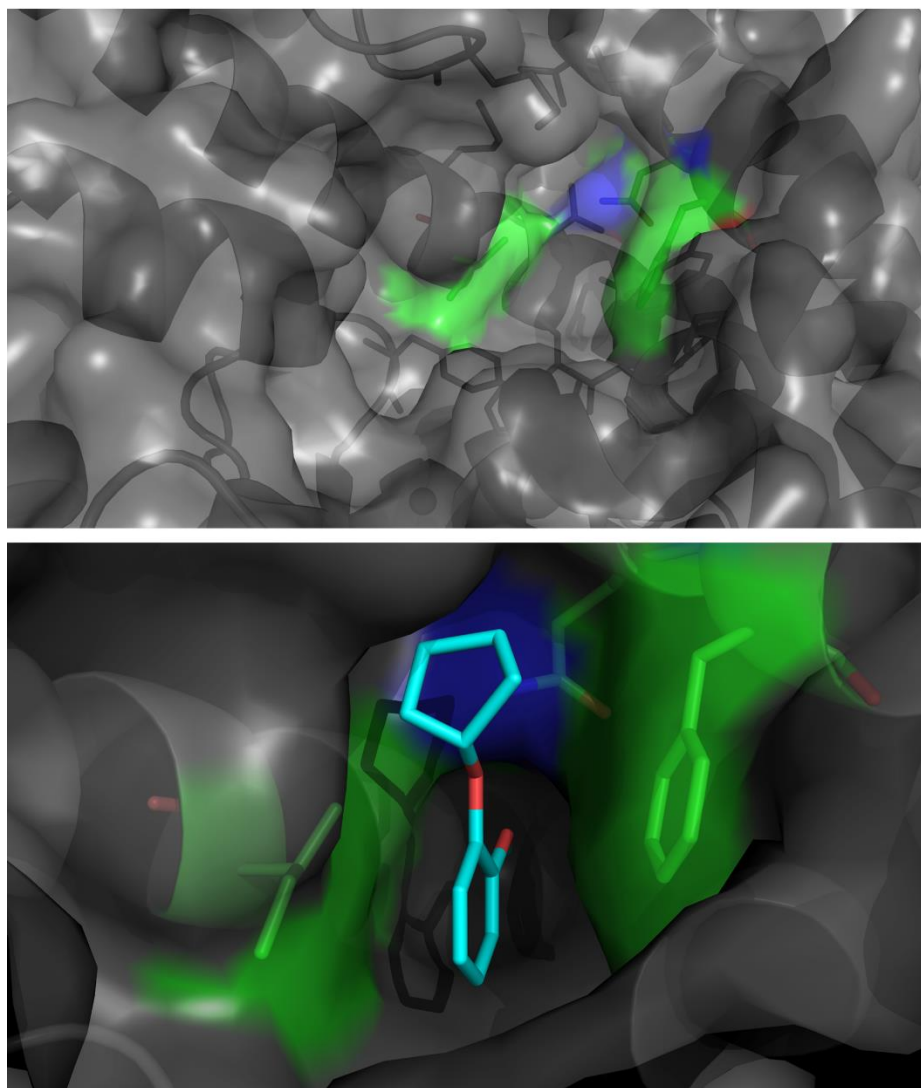


Figure 14: PDE4D catalytic domain with details of the residues involved in the ligand anchoring. Phe538 and Ile502 for P-clamp (green surface). Gln535 for Q-switch bind (blue surface). Catecholic core of an inhibitor (cyan sticks). PDB ID: 6F8T, adapted from Prosdocimi et al. 2018. Biochemistry 57 2876-2888

The last substituent on the phenyl is an extremely variable portion that in the majority of compounds is able to explore the M-pocket, in the proximity of the Zn²⁺ ion. Interestingly,

Rolipram displays two slightly different conformational poses for the M-pocket-exploring group. It is now believed that these two conformations reflect the ability of the inhibitor to orient itself toward the UCR2 domain, generating the HARBS/LARBS phenomenon.

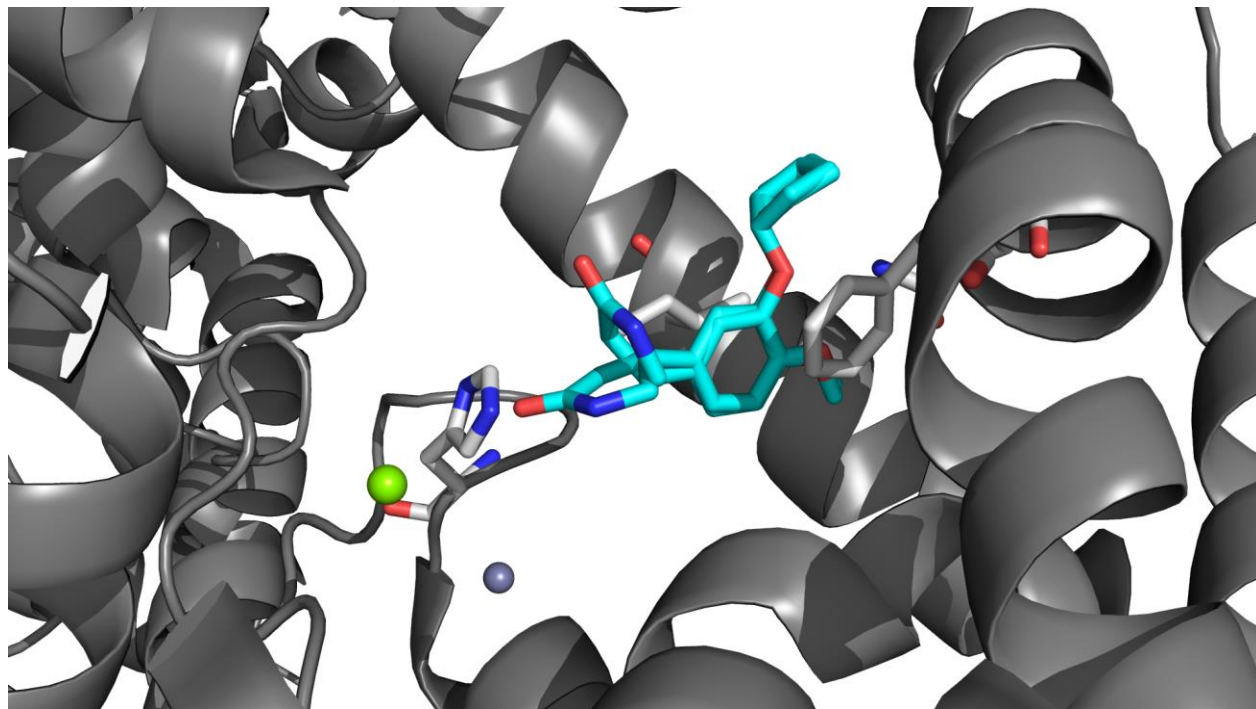


Figure 15: details of the two conformations that Rolipram can assume in PDE4 catalytic domain. PDB ID:3G4K. In the lower conformation, the M-pocket exploring group is oriented toward the metal ion coordination. In the higher conformation, the group look toward the externa part of the catalytic pocket.

This discovery, together with the functional studies of HARBS and LARBS^{47,77}, paved the way to the development of the modern inhibitors that show specificity for long isoforms and/or allosteric modulation.

Inhibitor Name	Structure	PDE1B	PDE2A	PDE3B	PDE4B	PDE4D	PDE5A	PDE7B	PDE8A	PDE9A	PDE10A	PDE11A
Zardaverine		>200	>200	1.5	0.93	0.39	81	>200	160	>200	14	140
(R,S)-Rolipram		>200	>200	>200	0.57	1.1	>200	>200	>200	>200	140	>200
Filaminast		190	89	9.4	0.96	1.0	53	0.99	120	>200	27	57
(R,S)-Mesopram		>200	>200	>200	0.42	1.1	>200	>200	>200	>200	63	>200
Cilomilast		87	160	87	0.025	0.011	53	44	7	>200	73	21
Roflumilast		>200	>200	>200	0.00084	0.00068	17	>200	>200	>200	>200	25
Piclamilast		68	54	11	0.000041	0.000021	3.5	8.8	>200	>200	21	1.6
Sildenafil		1.5	35	15	20	14	0.0022	78	>200	5.6	6.8	6.1
Vardenafil		0.3	3.1	0.58	3.8	3.9	0.0010	1.9	57	0.68	0.88	0.24
Tadalafil		50	130	280	9.2	19	0.0012	74	>200	150	19	0.010

The enzymes used are the catalytic domain of human PDEs. The numbers shown in the table are the 50% inhibition concentration (IC50) in μM .

Figure 16: Gold-standard PDEs inhibitors. From: Graeme L. Card et al. *Structure* 2004, vol 12, 2233-2247.

1.5.2 Allosteric modulation

One of the most important breakthroughs in PDE4 inhibition for cognition improvements is represented by the works of Gurney, Burgin and colleagues^{31,48,49}. Their pioneering research started from the discovery that the UCR2 regulatory domain is able to cap the catalytic active site of PDE4 in long isoforms (such as PDE4D3), generating an extended pocket with new druggable surfaces (see paragraph 1.3.7.).

The rationale behind their studies is threefold:

- PDE4 long isoforms are considered the best target for cognition improvements (see also paragraph 1.4.4). Exploiting the interactions with the capping domain UCR2, which can act as a lid only in the long isoforms, could improve the specificity of the inhibition, generating a molecule that may become (partially) sub-isoform selective.
- PDE4D UCR2 carries a key residue, Phe196 that is non conserved (Tyr274 in PDE4B) and that could therefore be exploited to induce isoform specificity.
- Since all highly potent PDE4 selective inhibitors display severe side effects, one possible strategy to overcome this problem could be the development of non-full inhibitors that attenuate emesis while remaining in the therapeutic window for cognitive improvements. This idea has already been tested in other drug development studies³¹.

Burgin and colleagues focused on the development of PDE4D inhibitors that, according to the rationale principles, could:

- link Gln535 (Q-switch, see paragraph 1.5.1 and Figure 17) to Phe196, stabilizing the closed and dimeric conformation of the enzyme.
- display a certain grade of specificity between the B and the D isoform, thanks to a core able to clamp Phe196 (PDE4D) better than Tyr274 (PDE4B9)

- display a decreased level of affinity for the P-clamp inside the pocket, generating a molecule that is not able to fully inhibit the enzyme.

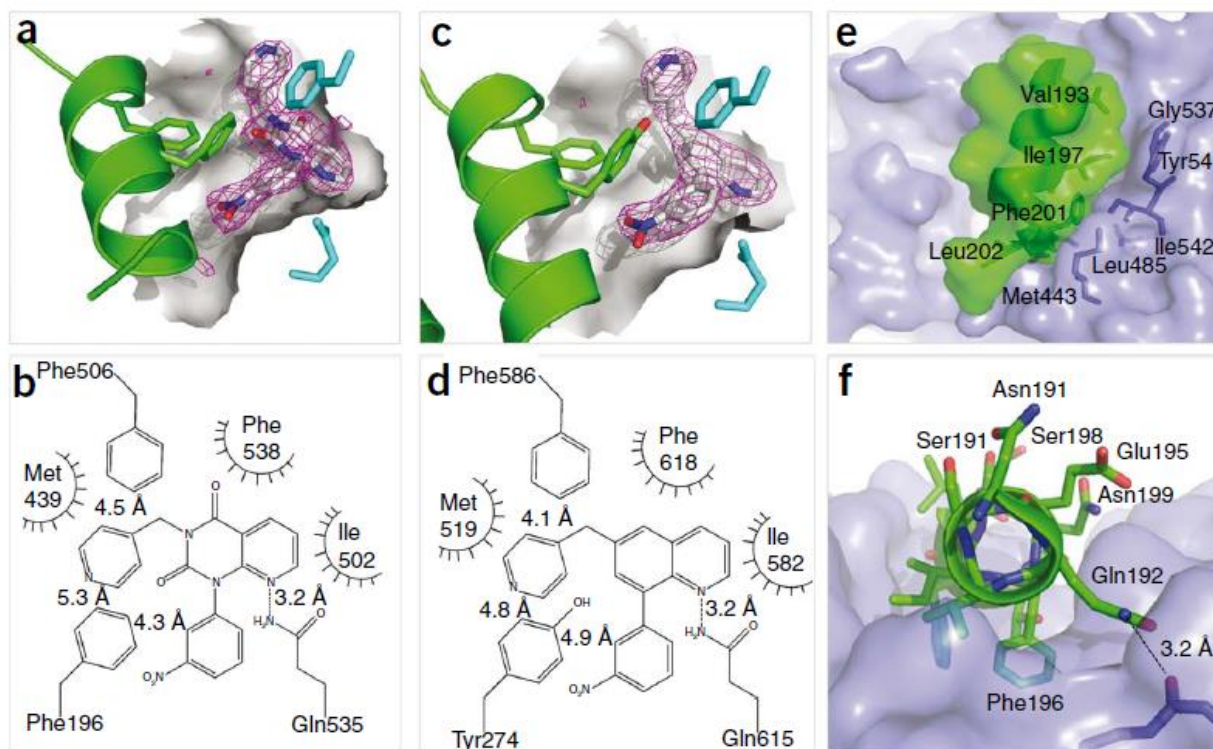


Figure 17: a, b) UCR2 capping and binding with a PDE4D allosteric modulator: structure (PDB ID: 3G4G) and scheme. c, d) UCR2 of a PDE4B capping and binding with a PDE4D allosteric modulator (PDB ID: 3G45). e, f) positioning of the UCR2 domain over the entrance to the catalytic pocket (PDB ID: 3G4G). Adapted from Burgin et al, Nature Biotechnology 2009.

Eventually, the molecules developed by Burgin and colleagues turned out to be negative allosteric modulators (NAMs). In fact, the capping of the catalytic pocket occurs in *trans*, which means that the UCR2 helix of a monomer closes the entrance to the pocket of the other monomer (see paragraph 1.3.4). The stabilization of the closed UCR2 conformation (by means of this kind of inhibitors) produces a stabilization of the whole dimer, decreasing the activity of the other catalytic unit³¹.

Following works by Burgin and colleagues allowed a better comprehension of PDE4-NAM, addressing the possibility of designing PDE4B-selective NAM⁴⁸.

As mentioned before, PDE4-NAMs retain potent benefits for cognition processes regulated by the cAMP-CREB pathway and are now been studied in highly promising clinical trials for XFS (Phase II) and Alzheimer's (Phase I).

1.6 The GEBR library

1.6.1 Historical overview

We refer to the “GEBR” library (GEnova BRuno) as a panel of catechol-based compounds for the inhibition of PDE4, developed by the medicinal chemistry group of Prof. Olga Bruno, at the University of Genova (Italy).

In 2004, the Bruno group reported their first panel of compounds⁷⁸. Since then, the library has constantly been improved and expanded with different kinds of chemical modifications. The main steps in the development of these compounds were:

1.6.1.1 2004 – First panel of inhibitors:

Synthesis and biological evaluation of several Rolipram-related compounds (3-cyclopentyloxy-4-methoxybenzaldehyde and 3-cyclopentyloxy-4-methoxybenzoic acid derivatives)⁷⁸. Some compounds showed inhibition of neutrophil activation. Among them, the three most active displayed the ability to increase cAMP levels in TNF α -stimulated neutrophils and one compound (now referred to as GEBR-4) showed PDE4 inhibition.

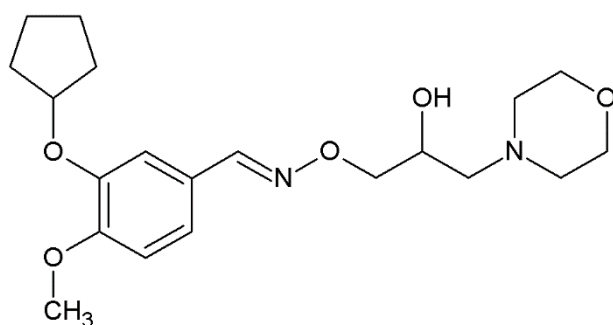


Figure 18: GEBR-4, the first promising GEBR compound

1.6.1.2 2009 – A new set of compounds display a slightly different activity on PDE4 isoforms of different length.

By 2009, the differential inhibition of PDEs had already been identified as crucial to reach a certain therapeutic window. Bruno et al. reported an expansion of the GEBR library and the evaluation of

the corresponding compounds for the inhibition of different PDEs²⁶. Among this subset, three compounds (GEBR-4a, GEBR-7b and GEBR-5b) showed a higher inhibition of PDE4D3 with respect to PDE4A4, PDE4B2 and PDE4C3.

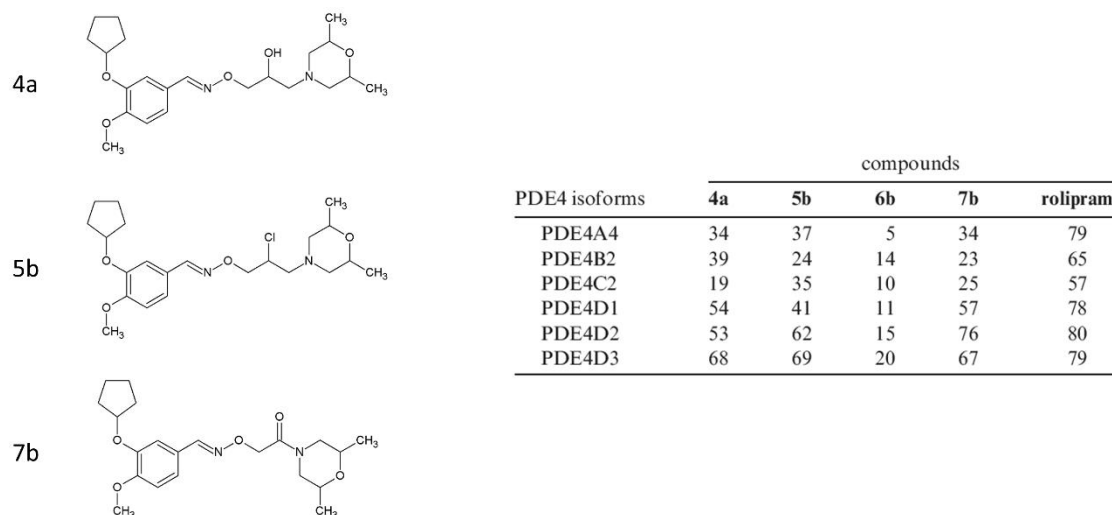


Figure 19: Structure and differential inhibition percentage of GEBR-4a, GEBR-5b and GEBR-7b. Table from Bruno et al. *J. Med. Chem.* 2009. 52.

1.6.1.3 2011 – First lead compound.

GEBR-7b is recognized as the best inhibitor within the GEBR library and Bruno and colleagues reported its effects in animal models⁷². The compound was shown not only to increase hippocampal cAMP levels and improve the late-phase consolidation process of spatial and object recognition in

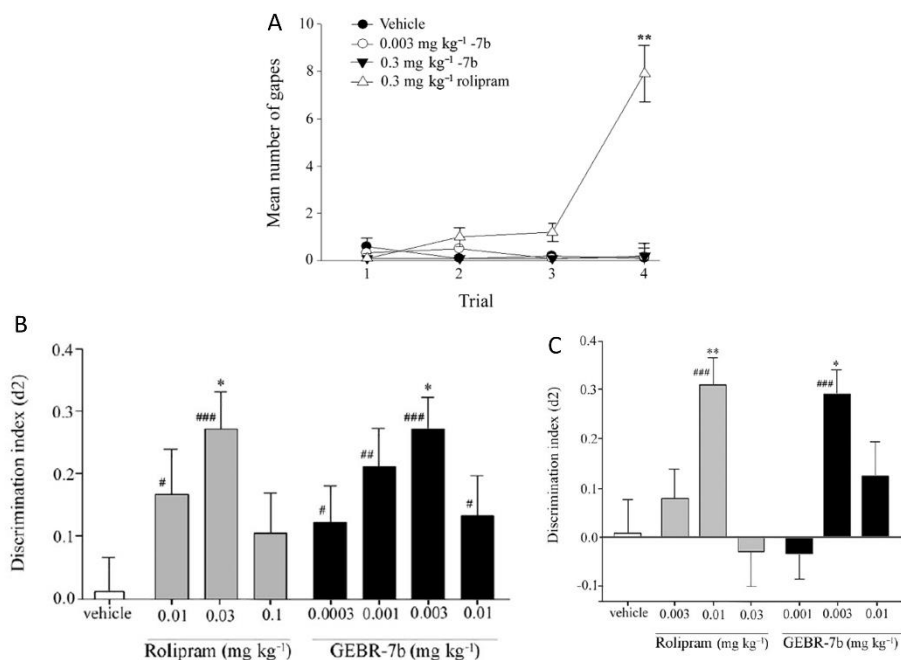


Figure 20: A) Influence of GEBR-7b and Rolipram in conditioned gaping tests for emesis evaluation. B) Influence of GEBR-7b and Rolipram on the discrimination index (d2) in the object location test. C) Influence of GEBR-7b and Rolipram on the discrimination index (d2) in the object recognition tests. Adapted from Bruno et al. *British Journal of Pharmacology.* 2011. 164. 2054-2063.

rats, but also to have lower emetic effects. Indeed, GEBR-7b showed memory consolidation at doses 3 to 10 times lower than Rolipram and dramatically reduced emetic effects up to a dose 100 times higher than the therapeutic one.

1.6.1.4 2014-2016 – Expansion of the library with GEBR-7b analogues.

In this period, the GEBR library was expanded considerably, often introducing modifications in the compounds, with particular attention to the linker connecting the catechol to the final morpholine^{79–81}. Moreover, variations were introduced also in the nature of the final morpholine, in its substitution with different moieties such as other heterocycles or bidentate polar tails, and in the fluorination of the catecholic methyl-ether. The works were always complemented with biological evaluations and computational studies (docking and molecular dynamics simulations) and yielded a considerable amount of promising compounds showing a certain degree of specificity for PDE4D3. One of the most promising inhibitor was GEBR-54 (previously called GEBR-8a or AR54), which was subsequently evolved into GEBR-32a (the current lead compound of the library) upon double fluorination of the catecholic methyl-ether. However, at that point in time, despite the improvements of the library, the details about the structural behavior of the compounds were still elusive.

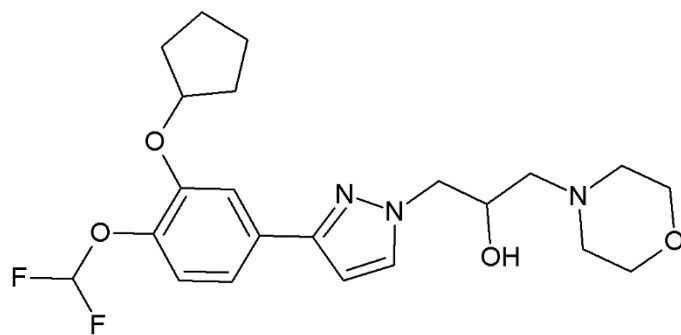


Figure 21: GEBR-32a

1.6.1.5 2017 – New lead compound, GEBR-32a.

In 2017, Ricciarelli *et al.* published a thorough and extremely promising report about the effects of GEBR-32a in rodents⁸². Indeed, GEBR-32a was tested *in vitro* and *in vivo* in mice models and, herein, the results are reviewed in detail:

- GEBR-32a triggers a significant increase in cAMP levels, both in cultured cells and hippocampal sections.
- preliminary *in vitro* toxicological investigations demonstrate that GEBR-32a is not cytotoxic or genotoxic, also at high concentrations (100 μ M)
- Object Location Task (OLT) was used to investigate episodic-like spatial memory, by evaluating if mice are able to recognize whether one of two different objects is moved from a position that has been previously experienced in a learning trial. Adult mice (vehicle treated) did not remember the object spatial arrangement after 24h from the learning trial. Mice treated with GEBR-32a in low doses (0.003 mg/kg) were able to perceive a shift in the object position and they spent considerably more time in the exploration of the new arrangement in the test trial. The same experiments were conducted on aged Tg2576 mice. Tg2576 mice are widely used as a model of AD. They overexpress a mutant APP (isof. 695, carrying Swedish mutation KM670/671NL) resulting in increased levels of A β and amyloid plaques⁸³. While WT aged mice show a normal functioning of the short-term memory (test trial is conducted 3 hour after the learning trial), Tg257 aged mice show an impaired

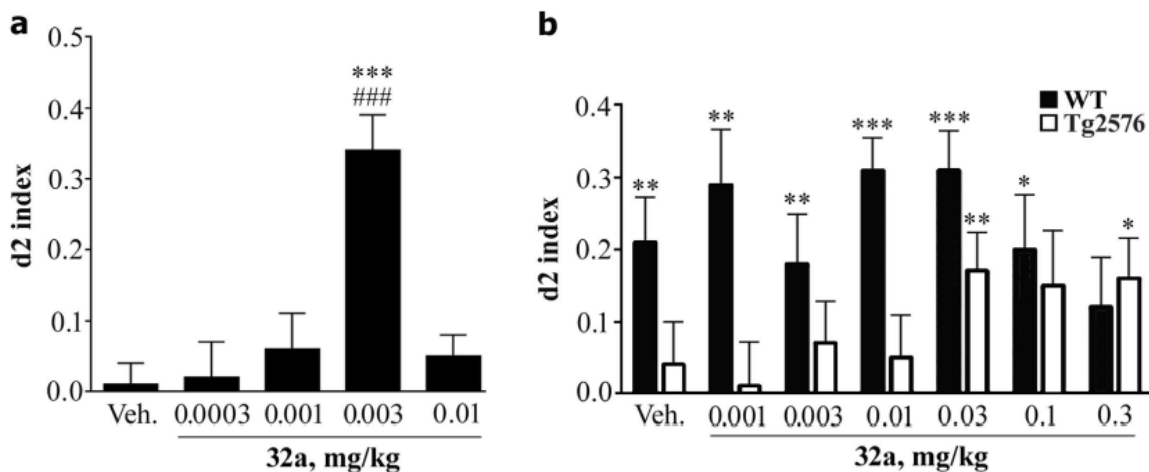


Figure 22. From Ricciarelli et al. *Scientific Reports* 2017. OLT tests. d2 index over 0 is considered positive. a) GEBR-32a administered to healthy adult mice 3 hours after the learning trial. b) GEBR-32a administered to aged WT and Tg2576 mice

functioning. Tg2576 treated with GEBR-32a did not recover the normal ability because the impairment is probably too severe. However, an acute administration of GEBR-32a triggered an improvement of short-term memory.

- Y-maze continuous task is used to assess the working memory. It measures the number of entries of mice into the three different arms of the maze. 50% of alternation is considered a chance level and reflects the absence of working memory. GEBR-32a administration did not improve the normal ability of adult mice (alternation level already over 50%). WT aged and Tg2576 mice showed memory impairment (alternation equal to chance level) and the first one displayed an improvement after GEBR-32a treatment. When chronically administered, GEBR-32a triggered an improvement also in Tg2576 mice.
- GEBR-32a was shown to be non-emetic in rats up to doses 100-1000 times higher than the precognitive ones.

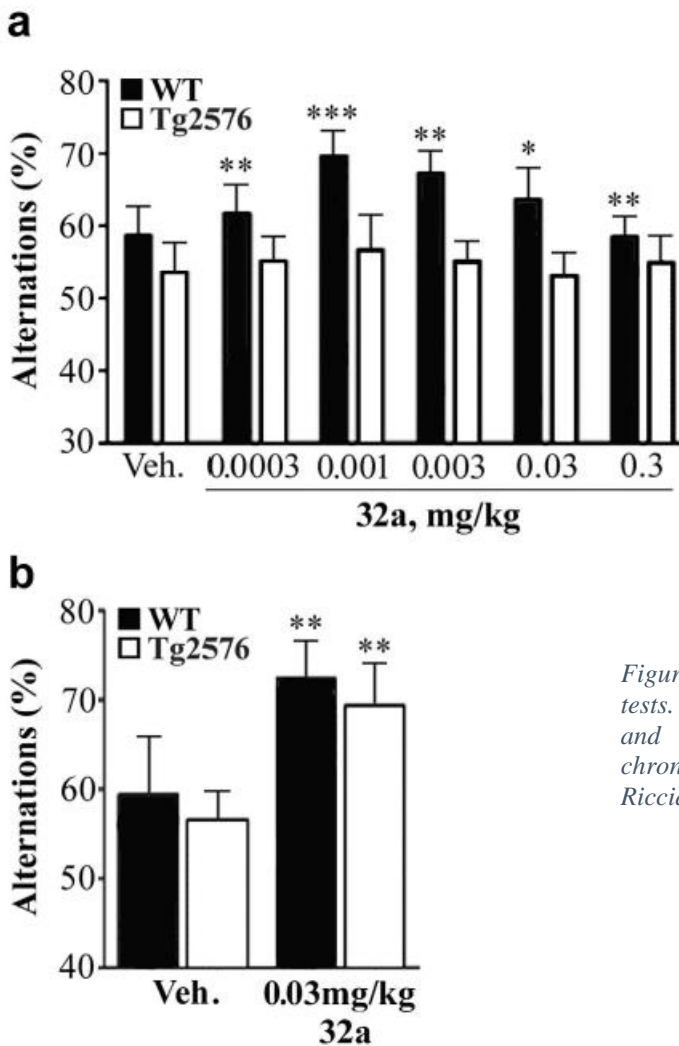
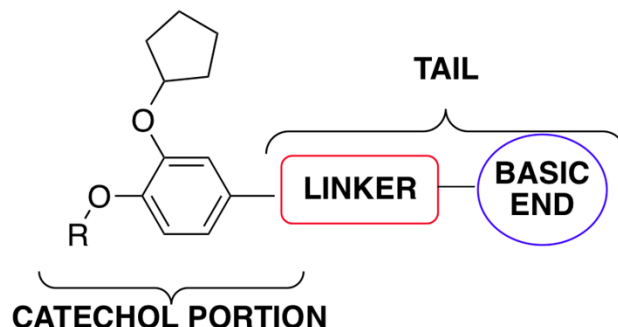


Figure 23. Percentage of alternation in Y-maze tests. a) Acute administration of GEBR-32a on WT and Tg2576 mice. b) GEBR-32a administered chronically in WT and Tg2576 mice. Adapted from Ricciarelli et al. Scientific Reports 2017.

1.6.1.6 2015 - 2018 – Structural studies

The structural bases of PDE4 inhibition by GEBR library compounds were experimentally-determined for the first time as the main PhD project described in this thesis⁸⁴.

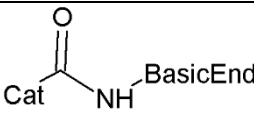
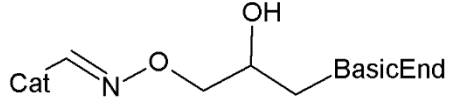
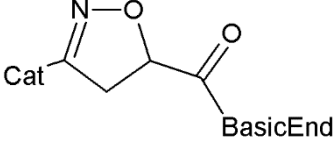
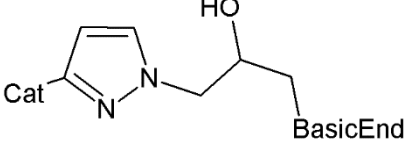
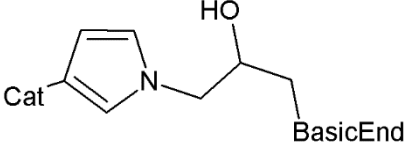
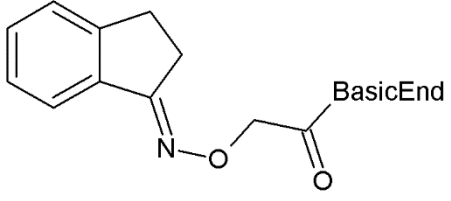
1.6.2 Chemical bases of the library



The GEBR library consists of compounds that feature a common catecholic scaffold. In these molecules, we can identify four main portions, three of which have been modified during the development of the library:

- Catecholic portion (invariant): Since the development of Rolipram⁶⁷, the catecholic moiety has been one of the most investigated scaffold in PDE4 drug development studies (see paragraph 1.5.1). All the compounds belonging to the GEBR library feature a common catecholic scaffold, with the substituent directed to the Q₂ pocket consisting of a cyclopentyl group.
- R group: This group is the Q₁-directed substituent to the catechol. The GEBR library features two different alternatives: a standard methyl or a difluoromethyl, the latter being introduced with the expansion of the library in 2015⁸¹.
- Linker: This portion is the most variable of the library. The linker has been modified in its chemical nature in order to explore different functionalities associated with the possible exploration of the M-pocket. Six families of linkers have been developed within the GEBR library, each comprising up to 10 members featuring alternative chemical modifications.

- Basic end: Terminal portion of the compounds. In 2009, Bruno *et al.* developed some GEBR-library compounds that are capable of inhibiting long and short isoforms differentially²⁶ (see paragraph 1.6.1.2). All the more active compounds that were also showing a differential behavior were featuring a morpholine or dimethylmorpholine as the terminal portion. Then, the GEBR library has been expanded with the para-hydroxy-piperidine terminal moiety, or as in the case of the subfamily of GEBR-26, with bidentate tails bearing two terminal hydroxyl groups (Figure 24).

Type	General Scaffold
Short linker	
Linear linker	
Isoxazoline-based linker	
Pyrazole-based linker	
Pyrrole-based linker	
Condensed cycle-based linker	

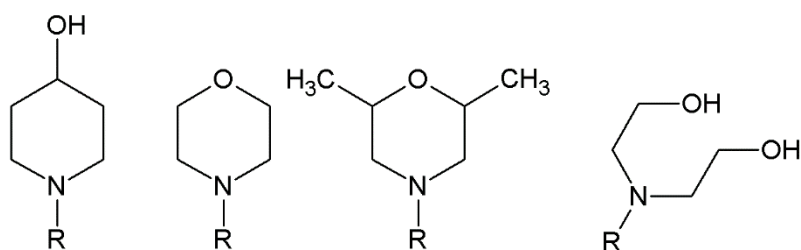


Figure 24: Different terminal portions in GEHR library compounds

1.7 Aim and overview of the PhD project: a mechanistic understanding of PDE4D inhibition by GEHR-library compounds

1.7.1 Origins

Studies exploring the therapeutic potentials of the GEHR library compounds have always provided extremely encouraging results, both *in vitro* and *in vivo*, for memory-enhancing and pro-cognitive effects. Together with PDE4-NAMs (which are currently the most promising inhibitors being developed), these compounds are now considered as an interesting strategy to overcome the severe side effects associated with PDE4 inhibition. However, a poor understanding of the molecular events that take place in the recognition between PDE4 and the GEHR-library compounds has always impaired further development of the library. Indeed, a mechanistic (and experimentally-driven) investigation of the GEHR library compounds was needed^{79,82}, in order to gain structure-function relationship data and pave the way for a structure-guided development of the next generation of compounds.

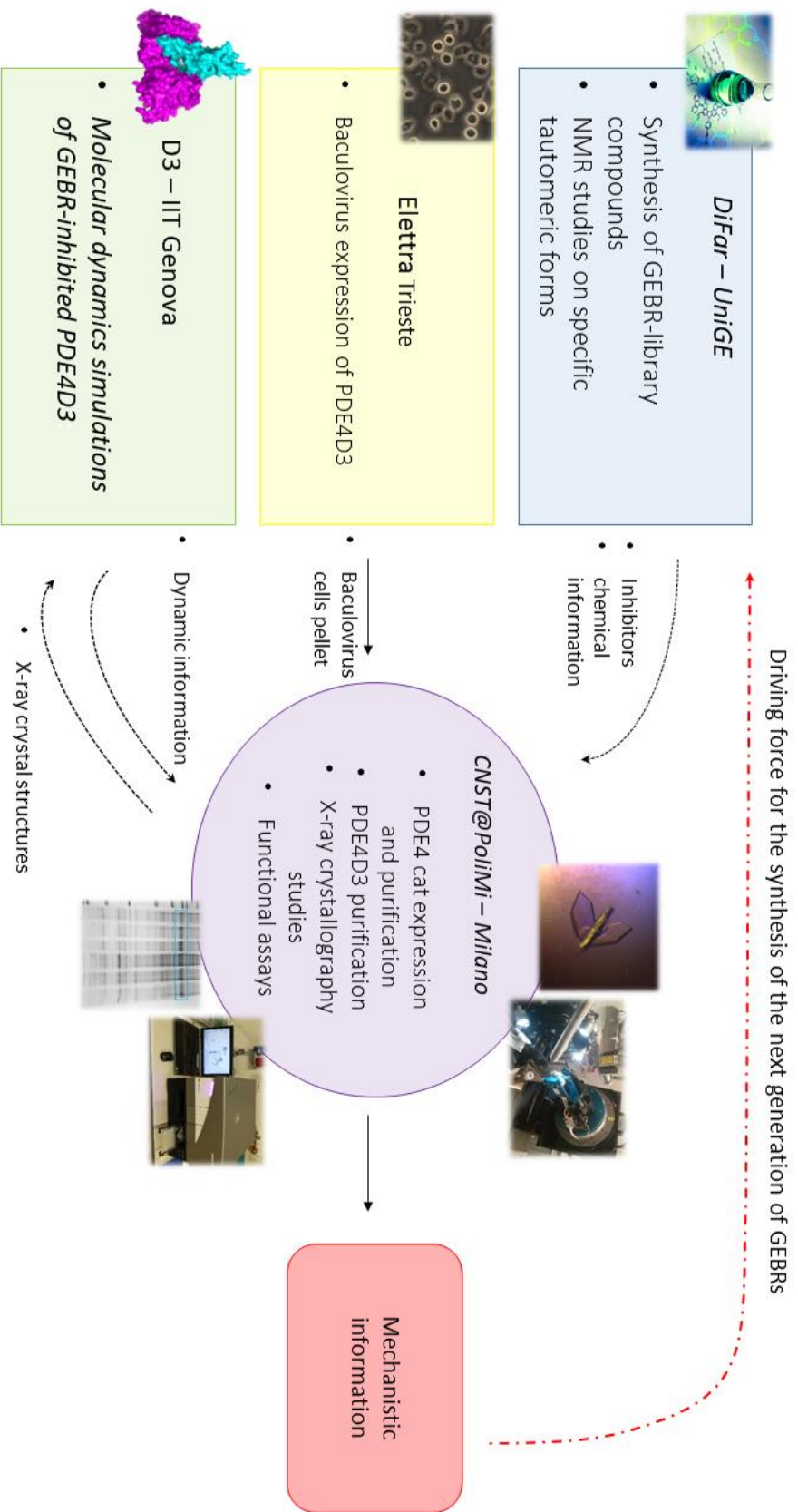
In 2015, a collaboration between the Bruno's group (DiFar, UniGe) and the Parisini's group (CNST@PoliMi, Istituto Italiano di Tecnologia) was established with the aim of characterizing the structural bases of PDE4 inhibition by the GEHR-library compounds, using X-ray crystallography studies of the inhibitors in complex with the PDE4D catalytic domain.

1.7.2 Expanding the project

In 2017, our first crystallographic structures were showing the existence of conformationally different groups of molecules. Encouraged by these results, the project was widened with the purpose of deepening the investigation from a functional point of view and expanding our knowledge on

biochemical behavior of the GEBRs. Indeed, the selected inhibitors of the library that were studied by X-ray crystallography were also analyzed with functional assays. The experiments were conducted on both the full length PDE4D3 and PDE4 catalytic domain , in order to complement and explain the insights derived by the X-ray structures, and to assess whether any structural difference in the binding pose of each inhibitor could have functional repercussions.

The capping of the catalytic domain by the UCR2 domain represents an extremely interesting regulatory mechanism, especially for drug development (see paragraph 1.3.4). Our mechanistic studies were not containing information about possible interactions and reciprocal behavior of GEBR compounds and the UCR2 domain. In fact, all the structural studies about PDE inhibition have been carried out using the catalytic domain only. Indeed, owing to their flexible regulatory domains, PDE4 long isoform are almost impossible to crystalize, and the most extended structure of a PDE4 long isoform (which is not complete) has been obtained by an engineered stabilization of the regulatory portions ²¹. To gain some information in this direction, the project in 2017 was further expanded, introducing Molecular Dynamics simulation of the best available PDE4 model, carrying the GEBR compounds docked inside the catalytic pocket in our experimental pose. With this approach, we aimed at developing an understanding of the dynamic phenomena associated with PDE4 inhibition, in a structural context where the UCR2 domain lays in its closed conformation.



2 Materials and methods

2.1 Phosphodiesterase 4D catalytic domain expression and purification

2.1.1 Construct

The pET15b vector (Novagen) containing the cDNA encoding human full length PDE4D3 fused at its C-terminus with a 6xHis tag was purchased from Eurofins MWG.

The DNA fragment encoding the catalytic domain (amino acid range: 244-578) was amplified via polymerase chain reaction (PCR) from the full length PDE4D3 gene and cloned by conventional methods into a pET3a vector (Novagen). The C-terminal His-tag DNA sequence was added during PCR amplification.

Primers used for PDE4 catalytic domain amplification:

- Forward primer: ttttcatatgtccattccgcggtttggg
- Reverse primer: cccccggatcctcagtgatgatgatgatgatgtgctgcgggat

The final construct was verified upon sequencing (sequencing outsourced to myGATC).

```
      1      10      20      30      40      50      60
PDE4D3 MMHVNNFPFRRHSWICFDVDNGTSAGRSPLDPMTSPGSLILQANFVHSQRRESFLYRSD
PDE4D  .....
```



```
      70      80      90      100     110     120
PDE4D3 SDYDLSPKMSMRNSSIASDIHGDDLIVTPFAQVLASLRTVRNNFAALTNLQDRAPSKRSP
PDE4D  .....
```



```
     130     140     150     160     170     180
PDE4D3 MCNQPSINKATITEEAYQKLASETLEELDWCCLDQLETLQTRHSVSEMASNKFKRMLNREL
PDE4D  .....
```



```
     190     200     210     220     230     240
PDE4D3 THLSEMSRSGNQVSEFISNTFLDKQHEVEIPSPQTQKEKEKKRPMMSQISGVKKLMHSSSL
PDE4D  .....
```



```
     250     260     270     280     290     300
PDE4D3 TNS SIPRFGVKTEQEDVLAKELEDVNKWGLHVFRIAELSGNRPLTVIMHTIFQERDLLKT
PDE4D  ..M SIPRFGVKTEQEDVLAKELEDVNKWGLHVFRIAELSGNRPLTVIMHTIFQERDLLKT
```



```
     310     320     330     340     350     360
PDE4D3 FKIPVDTLITYLMTLEDHYHADVAYHNNIHAADVQSTHVLLSTPALEAVFTDLEILAAI
PDE4D  FKIPVDTLITYLMTLEDHYHADVAYHNNIHAADVQSTHVLLSTPALEAVFTDLEILAAI
```

```

          370      380      390      400      410      420
PDE4D3 FASA IHDVDHPGVSNQFLINTNSELALMYNDSSVLENHHLAVGFKLLQEENC DIFQNLTK
PDE4D  FASA IHDVDHPGVSNQFLINTNSELALMYNDSSVLENHHLAVGFKLLQEENC DIFQNLTK

          430      440      450      460      470      480
PDE4D3 KQRQSLRKMVIDIVLATDMSKHMNLLADLKT MVETKKVTSSGVLLLDNYS DRIQVLQNMV
PDE4D  KQRQSLRKMVIDIVLATDMSKHMNLLADLKT MVETKKVTSSGVLLLDNYS DRIQVLQNMV

          490      500      510      520      530      540
PDE4D3 HCADLSNPTKPLQLYRQWTD RIMEEFFRQGDRE RERERGMEISPMCDKHNASVEKSQVGFID
PDE4D  HCADLSNPTKPLQLYRQWTD RIMEEFFRQGDRE RERERGMEISPMCDKHNASVEKSQVGFID

          550      560      570      580      590      600
PDE4D3 YIVHPLWETWADLVHPDAQD ILDTLEDNREWYQSTIPQSPSPAPDDPEEGRQGQTEKFQF
PDE4D  YIVHPLWETWADLVHPDAQD ILDTLEDNREWYQSTIPQAHHHHHH.....

          610      620      630      640      650      660
PDE4D3 ELTLEEDGESDTEKDSGSQVEEDTSCSDSKTLCTQDSESTEIPLDEQVEEEEAVGEEEEESQ
PDE4D  .....

          670
PDE4D3 PEACVIDDRSPDT
PDE4D  .....

```

Figure 25: Alignment between PDE4D3 gene and PDE4D catalytic domain construct

2.1.2 Protein Expression

The pET3a-PDE4 (catalytic domain) plasmid was inserted into *E.coli* BL21(DE3) pLysS cells (Invitrogen) for overexpression. Then, cells were cultured at 37° in Luria-Bertani (LB) broth supplemented with 50 mg/L ampicillin, until OD₆₀₀ reached 0.6.

In order to optimize the production yield, test-expression were performed in small scale cultures (0.1 L) varying inductor concentration, temperature and induction period. After yield optimization, several production batches (4 L each) were necessary to fulfill the demand for the final protein

sample. Expression was induced with 1-thio- β -D-galactopyranoside (IPTG, final concentration: 0.5 mM), and the culture was left on an orbital shaker at 25°C overnight.

Detailed protocol:

1. Add 1-4 μ L (100 ng circa) of the pET3a-PDE4 (catalytic domain) vector to a one-shot vial (50 μ L) of *E.coli* BL21(DE3) cells (50 μ L)
2. Incubate on ice for 15 minutes
3. Heat-shock at 42°C for 45 seconds
4. Incubate on ice for 2 minutes
5. Add 1ml LB
6. Incubate at 37°C for 1h at 180 rpm
7. Plate cells on LB-agar (ampicillin 50 mg/L)
8. Incubate at 37°C overnight
9. Select a colony and prepare 40 ml (10 ml for each liter of culture) of pre-inoculum of LB-amp.
10. Incubate at 37°C overnight (shaking cultures at 180 rpm)
11. Use the pre-inoculum to start 4L of large scale culture.
12. Grow at 37°C, 180 rpm, until OD reaches 0.6.
13. Add IPTG to a final concentration of 0.4 mM
14. Incubate at 25°C overnight.

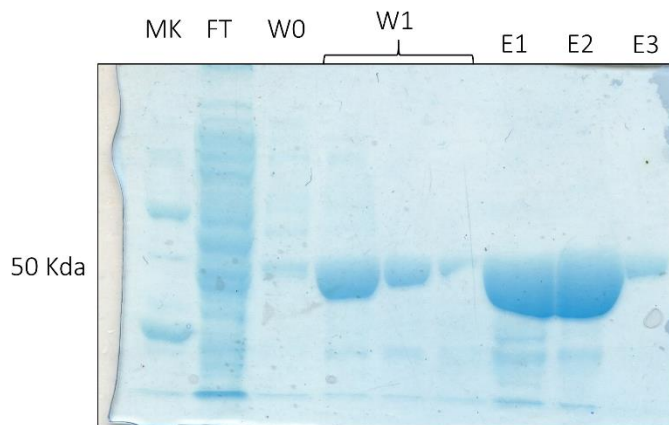
2.1.3 Purification

The PDE4D catalytic domain was purified following a three-step protocol involving Ni²⁺ affinity chromatography, size-exclusion chromatography and ion-exchange chromatography.

Detailed protocol:

1. After overnight expression, pellet cells at 10000 rpm for 10' at 4°C
2. Resuspend cells to a final concentration of 0.2 g/ml in lysis buffer (50 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol, PMSF 1mM, DNase 0.05U/ml, β -mercaptoethanol 100% 0.08 μ L/ml)
3. Disrupt cells by sonication (3' burst, 2' break – repeated 8-10 times – duty cycle 50 – output control 5)

4. Pellet cell debris by centrifugation at 18000 rpm for 45' at 4°C
5. Filter supernatant
6. Add supernatant into the chromatography gravity column, packed with 2ml of Ni-NTA resin. Collect the FT
7. Wash the resin with 5 column volumes of wash buffer 1 (50 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol)
8. Wash the resin with 10 column volumes of wash buffer 2 (50 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol, 40 mM Imidazole pH 7.5)
9. Elute in multiple fractions by elution buffer (50 mM Hepes pH 7.5, 150 mM NaCl, 5% glycerol, 400 mM Imidazole pH 7.5)

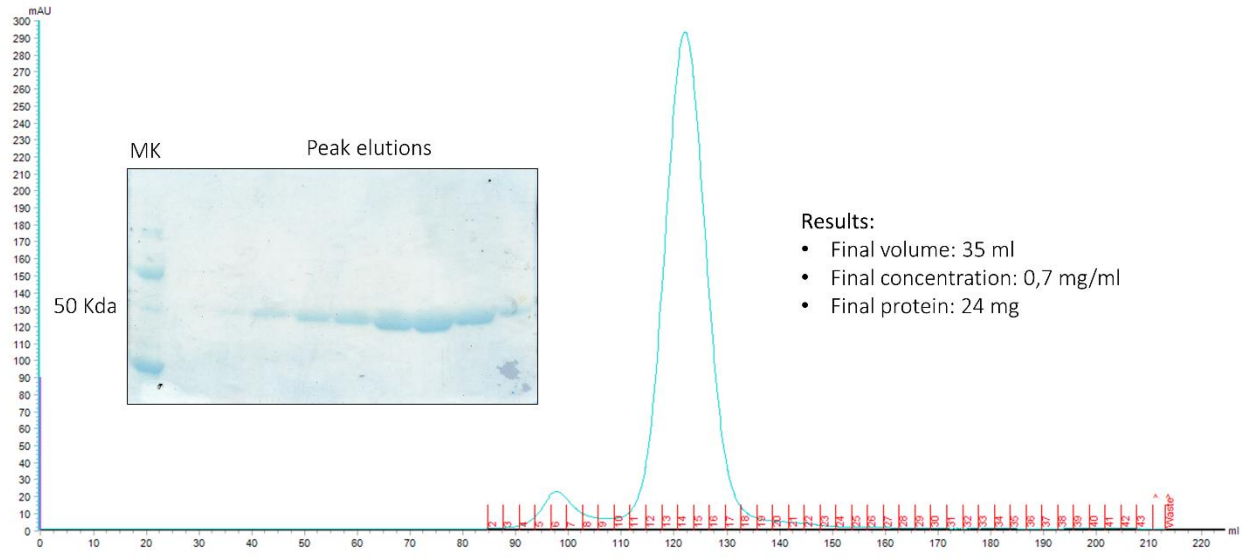


Results:

- Final volume: 4ml
- Final concentration: 7,5 mg/ml
- Final protein: 30 mg

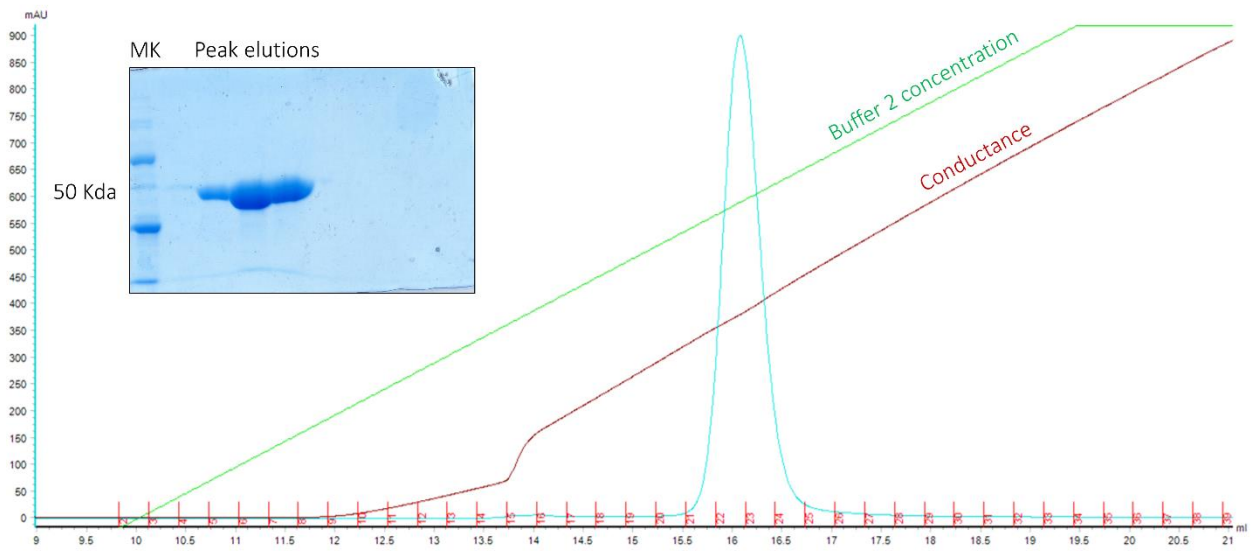
10. Size-exclusion chromatography:

- a. performed using a Hi-Prep 26/60 sephacryl s100 column (GE-Healthcare)
- b. Buffer: 50 mM HEPES pH 7.5



11. Ion-exchange chromatography:

- performed using a Hi-Trap Q HP column (GE-Healthcare)
- Buffer 1: 50 mM HEPES pH 7.5
- Buffer 2: 50 mM HEPES pH 7.5, 1M NaCl

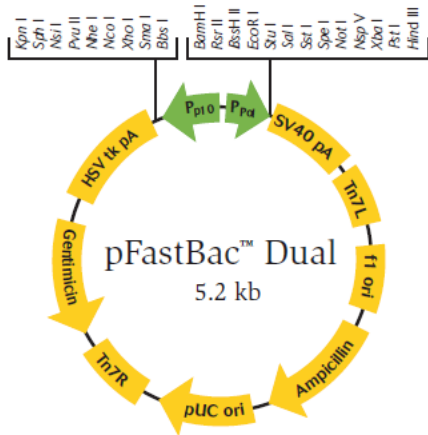


12. Change buffer by overnight dialysis to 50 mM HEPES pH 7.5 or 20 mM Tris pH 7.5

The final protein samples were divided into two main aliquots: 1) samples for crystal growth (5-10 mg/ml) and 2) sample for enzymatic assays (0.01-0.05 mg/ml). The protein samples were concentrated by Amicon Ultra centrifuge units (MWCO 10 kDa).

2.2 Phosphodiesterase 4D3 expression and purification

2.2.1 Construct



The human PDE4D3 gene with a C-terminal 6xHis-tag, codon optimized for insect cells (Sf9) and inserted into the pFastBac Dual vector between the BamHI and the NotI restriction sites, was purchased from GenScript. The construct carries the mutations Ser54Asp and Ser579Ala. The first one is introduced in order to mimic the activating phosphorylation by PKA, whereas the second one to avoid the inactivating mutation by ERK³¹.

Figure 26: From "Bac to Bac, baculovirus expression system" – Invitrogen.

In the following, the DNA and protein sequence of the PDE4D3 construct is reported. Highlighted are the phosphomimetic mutation (green), the phosphoabolishing mutation (red) and the histidine tag (yellow)

```

atgatgcacgtcaacaacttttttaggagacatagctggatctgctttgatgtggataat
M M H V N N F F R R H S W I C F D V D N
ggaaccaggcgcaggacgttccccctctcgatcccacatctccagggtccggccttata
G T S A G R S P L D P M T S P G S G L I
ctccaagcgaatctcgtgcattcgcagcgcagagaggatTTTTGTACAGGAGTGACAGC
L Q A N F V H S Q R R E D F L Y R S D S
gactatgacctgtcccccaaatcgatgtccaggaactcttcgatcgcgctccgatattcat
D Y D L S P K S M S R N S S I A S D I H
ggtgacgacctgatcgtgacgccttttgacaggtcctggcaagccttagaacggttaga
G D D L I V T P F A Q V L A S L R T V R
aataatttcgctgccttgacgaatctccaagatcgcgctccatcgaagcgtccccgatg
N N F A A L T N L Q D R A P S K R S P M
tgtaaccagcctagtattaataaagcaactatcaccgaagaagcttaccaaaagttggca
C N Q P S I N K A T I T E E A Y Q K L A
agcgaaacacttgaagagctggattggtgtctcgatcagttggagaccttgacagctagg
S E T L E E L D W C L D Q L E T L Q T R
catagcgtctccgaaatggccagcaacaagttaagcgcagttgaaatagagagttgaca
H S V S E M A S N K F K R M L N R E L T
cacctctcggaaatgtctcgtagcggaaaccaagtaagtgaatttatatcaaatccttt
H L S E M S R S G N Q V S E F I S N T F

```

ttggataagcagcagcagggctcgaaatccctctcccactcaaaaggagaaggaaaagaag
 L D K Q H E V E I P S P T Q K E K E K K
 aagcgccccatgagtc aaatttctggcgtaaagaagctgatgcatagtagcagctctcagc
 K R P M S Q I S G V K K L M H S S S L T
 aactcaagcattccacggttctcggtttaa aacggagcaggaagatggttctcgcaaaggaa
 N S S I P R F G V K T E Q E D V L A K E
 ctcgaaagacgtgaacaagtggggattgcatgtgttccgattgctgaactttccggcaat
 L E D V N K W G L H V F R I A E L S G N
 agacccttacgggtgataatgcacacgattttccaagagagagacctcctgaagaccttc
 R P L T V I M H T I F Q E R D L L K T F
 aaaatcccagtagacactttgatcacgtatcttatgaccttgaggaccactatcacgcc
 K I P V D T L I T Y L M T L E D H Y H A
 gatgtggcgtagcacaataacattcatgcggcagatgtgggtgcaaagcacgcatgtcttg
 D V A Y H N N I H A A D V V Q S T H V L
 ctgtctactcctgcactggaggcagttttcacccgatcttgaaatattggccgccatTTTT
 L S T P A L E A V F T D L E I L A A I F
 gcttccgctatccatgatgtggatccgggtgtgtccaaccagttcttgataaaatcc
 A S A I H D V D H P G V S N Q F L I N T
 aactcagaattggcactgatgtacaatgactcgtctgtcttgaaaaccaccacctggca
 N S E L A L M Y N D S S V L E N H H L A
 gtaggctttaaactgctgcaggaggagaactgcgatatattccaaaacctgacaaagaaa
 V G F K L L Q E E N C D I F Q N L T K K
 cagaggcagtcgcttaggaaaatggtaatcgacattgttcttgccacagacatgtctaaa
 Q R Q S L R K M V I D I V L A T D M S K
 catatgaatctgctggccgatcttaaaactatggttgaaactaaaaagggtcacatcgtct
 H M N L L A D L K T M V E T K K V T S S
 ggtgtattgctcctggataactattcggaccgcacccaggtactgcaaaataggtacat
 G V L L L D N Y S D R I Q V L Q N M V H
 tgccgagatttgagcaatcccaccaagcccctccagctgtatcgccaatggacagaccgc
 C A D L S N P T K P L Q L Y R Q W T D R
 atcatggaagaatttttctcgtcaaggtgatcgtgagagggagcgcggtatggagatctca
 I M E E F F R Q G D R E R E R G M E I S
 cctatgtgcgataaacataacgcctcagtcgaaaagtcccaggtcgggttttatagattat
 P M C D K H N A S V E K S Q V G F I D Y
 attgttcatccactgtgggaaactgtgggctgatttggttcaccaccgacgccagacatc
 I V H P L W E T W A D L V H P D A Q D I
 ttggatactctggaggataaccgcgaatggtagctactatcccacagcgccctca
 L D T L E D N R E W Y Q S T I P Q A P S
 cctgtcctcggacgatcctgaagaaggctcgtcagggccaaacagagaaaattccaattcgag
 P A P D D P E E G R Q G Q T E K F Q F E
 ctgactctggaagaagacgggtgagctgataccgagaaagactcgggctcgcaggtcgag
 L T L E E D G E S D T E K D S G S Q V E
 gaggacattcatgttcagattcgaagactctctgcacacaggatttcggaagcacagag
 E D T S C S D S K T L C T Q D S E S T E
 atacccttgacagcaagttgaggaagaagcgggtggcgaggaagaagagtcacaacc
 I P L D E Q V E E E A V G E E E E S Q P
 gaggcctgtgtgatcgatgataggtcgccctgatacaggccaccaccaccaccatcac
 E A C V I D D R S P D T G H H H H H H

2.2.2 Bacmid generation

For the generation of the recombinant bacmid, *E.coli* DH10Bac and *E.coli* EMBacY cells (kindly provided by Prof. I. Berger, University of Bristol, UK) were used. These strains carry:

- Baculoviral DNA with LacZ gene that contains mini Tn7 attachment site sequence for transposition. Successful integration of the gene of interest into viral DNA results in the

disruption of the LacZ alpha subunit, which leads to growth of colonies during IPTG/Bluo-gal selection.

- Helper plasmid with gene for transposase

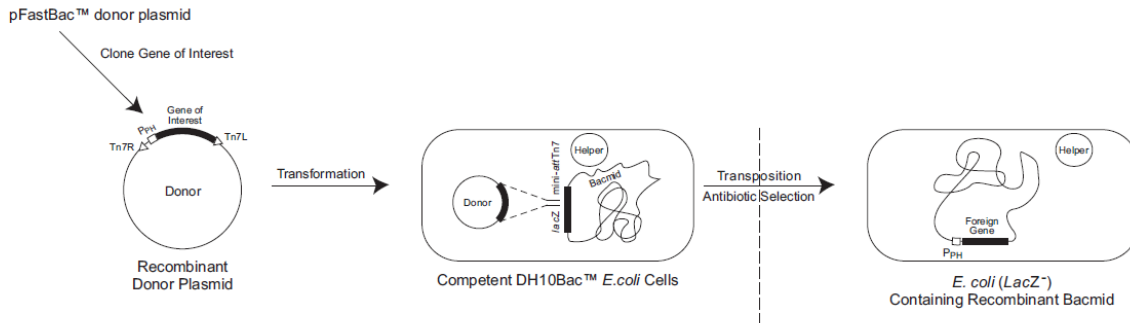


Figure 27: The mechanism for bacmid generation (adapted from the *Thermofisher bac-to-bac manual*)

The following procedure was used for bacmid generation:

1. Transform 100 µl of DH10Bac competent cells with 0.5 µl of pFastBacDual-PDE4D3 previously diluted 1:2 (250 ng of total DNA).
2. Incubate 30 minutes in ice.
3. Heat shock at 42°C for 45 seconds.
4. Incubate in ice for 2 minutes.
5. Add 900 µl of SOC medium (tryptone 1% w/v, yeast extract 0.5% w/v, NaCl 10 mM, KCl 2.5 mM, MgCl₂ 10 mM, glucose 20 mM)
6. Incubate 6-7 hours, 37°C, 600 rpm
7. Plate 50 µl of 1x, 10x and dilution of transformation mix on 3ml of LB-agar plates (6-wells plate) supplemented with:
 - a. Kanamicin 50 µg/ml (selection for resistance gene carried by the bacmid)
 - b. Tetracyclin 10 µg/ml (selection for resistance gene carried by the helper plasmid)
 - c. Gentamycin 7 µg/ml (selection for resistance gene carried by pFastBac Dual, the gene is integrated via transposition into the bacmid)
 - d. IPTG 40 µg/ml (blue-white screen)
 - e. Bluo-Gal 40 µg/ml (blue-white screen)

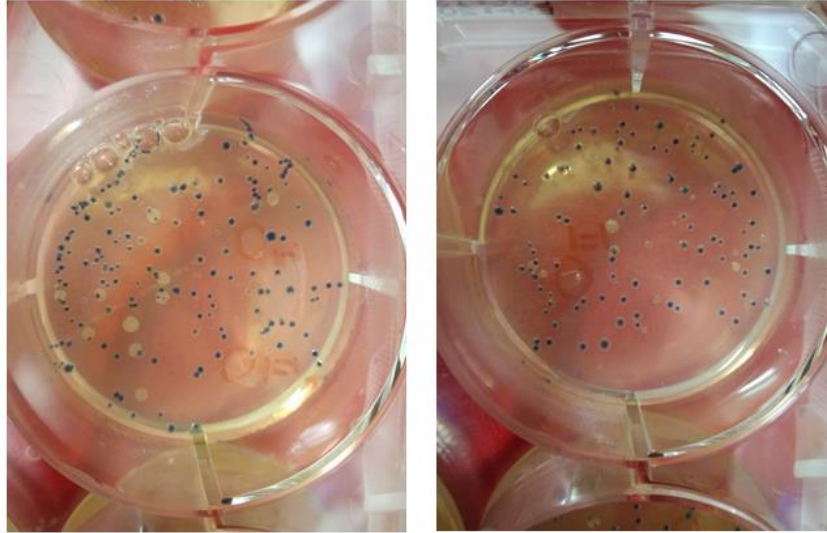


Figure 28: Bacmid generation for PDE4D3 construct. Transposition made by *E.coli* DH10Bac. Blue-white screen of *E.coli* DH10Bac transformed cells.

A white colony was isolated and grown overnight (37°, 180 rpm) in 3 ml LB supplemented with 50 µg/ml kanamycin and 7 µg/ml gentamycin. The bacmid was isolated using the following protocol:

1. Centrifuge cells at 5000 rpm for 10 minutes, 4°C, discard supernatant
2. Resuspend the cell pellet in 300 µl of buffer P1
3. Add 300 µl of buffer P2, inversion 5-7 times, incubation for 5 minutes at RT
4. Add 300 µl of buffer P3, inversion 5-7 times
5. Centrifuge at 13000 rpm for 10 minutes, RT, discard pellet
6. Transfer supernatant in a new tube
7. Add 700 µl of IPA 100%
8. Centrifuge at 13000 rpm for 30 minutes, 4°C
9. Discard carefully the supernatant by pipetting it
10. Wash DNA with 200 µl EtOH
11. Centrifuge at 13000 rpm for 5 minutes, RT
12. Discard carefully the supernatant under the laminar flow hood to keep the sample sterile.
13. Air-dry the DNA pellet under the laminar flow hood.
14. Resuspension in 50 µl sterile MilliQ water.

Buffer details for bacmid prep procedure:

- P1: resuspension buffer
 - 50mM Tris-Cl pH 8.0
 - 10mM EDTA
 - 100ug/mL RNase A
- P2: lysis buffer
 - 200mM NaOH
 - 1% SDS
- P3: neutralization buffer
 - 3.0 M potassium acetate pH 5.5

2.2.3 Generation of the virus – V0

The generation of the virus is carried out in Sf9 cells plated in a 6-well plate. Each experiment is done in duplicate with one well containing non transfected cells as a negative control. Procedure:

1. Plate 0.5×10^6 sf9 cells/ml in each well in a final volume of 3 ml.
2. Let the cells adhere at RT for 15 minutes.
3. Prepare transfection reaction:
 - a. In tube A add 100 μ l of medium ESF921 + 50 μ l of recombinant bacmid
 - b. In tube B add 100 μ l of medium ESF921 + 10 μ l FugeneHD transfection reagent (Promega)
 - c. Add tube B content into tube A and mix.
 - d. Incubate 5 minutes at RT.
4. Add 150 μ l of transfection reaction dropwise to each well
5. Incubate the cells 48-60 hours at 27°C
6. Check the cells daily for contamination and growth. Infected cells stop proliferating and appear slightly bigger than the control cells.
7. Collect exhausted medium containing virus by centrifugation at 1000 x g for 10 minutes and discard the cells. Virus is secreted from the cells.
8. Add FBS at a final concentration of 2% to stabilize the virus

2.2.4 Amplification of the virus – V1,2

High titer virus is obtained by two rounds of amplification.

V1 Preparation:

1. Seed 1.6×10^6 cells in 5 ml of ESF921 in T25 flask.
2. Add 150 μ l of V0 virus
3. Incubate 4 days at 27 °C, in adhesion
4. Collect exhausted medium containing virus by centrifuge at 1000 x g for 10 minutes, RT
5. Add FBS at a final concentration of 2% to stabilize the virus and store it at 4 °C.

V2 preparation:

1. Split Sf9 cells at density of 1×10^6 cells/ml in 50 ml of ESF921 in a 125 ml flask
2. Add 150 μ l of V1 virus.
3. Incubate 4-5 days at 27 °C, 130 rpm
4. Collect exhausted medium containing virus by centrifuge at 1000 x g for 10 minutes, RT
5. Add FBS at a final concentration of 2% to stabilize the virus and store it at 4 °C.

After the second cycle of amplification proceed with test expression.

2.2.5 Test expression

Test expression was done using Sf9 and Hi5 cells in a 24-well plate, following the procedure:

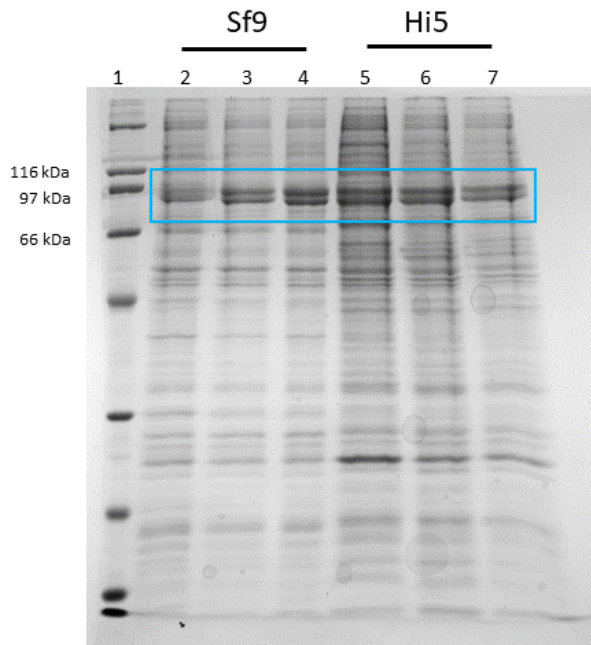
Expression condition:

- 3×10^6 cells/well at a final density of 1×10^6 cells/ml
- Tested 3 quantities of virus (MOIs):
 - **MOI1** (1 μ l of P2 virus amplification/ 3×10^6 cells)
 - **MOI2** (10 μ l of P2 virus amplification/ 3×10^6 cells)
 - **MOI3** (100 μ l of P2 virus amplification/ 3×10^6 cells)
- Incubate 3 days at 27 °C, 850 rpm

Purification:

- Centrifuge cells at 1000 x g for 10 minutes at 4 °C
- Resuspend the pellet in 1ml of lysis buffer (50mM Tris pH 8, 0,5M NaCl, 10mM imidazole, 10mM MgCl₂, 1x complete protease inhibitor cocktail (Roche), 10U of benzonase)
- Lyse cells by 2 cycles of freeze and thaw

- Centrifuge at 16000 x g for 10 minutes, 4°C
- Add 20µl of Ni-NTA agarose resin (Qiagen) into each supernatant sample and incubate for 1h at 4°C in agitation
- Wash the resin with 2 ml of wash buffer (50mM Tris pH 8, 0,5M NaCl, 10mM imidazole)
- Elute with 50µl of elution buffer (50mM Tris pH 8, 0,5M NaCl, 300mM imidazole)



1. Marker, Biorad ladder, 5µl
2. Sf9 E MOI 1, 10µl
3. Sf9 E MOI 2, 10µl
4. Sf9 E MOI 3, 10µl
5. Hi5 E MOI1, 10µl
6. Hi5 E MOI2, 10µl
7. Hi5 E MOI3, 10µl

Figure 29: PDE4D3 test expression in insect cells

Based on the results of the expression tests, Sf9 cells in combination with MOI3 of V2 virus were chosen for the scale-up conditions as they yielded samples of higher purity after Ni-NTA affinity chromatography purification.

Some protein degradation problems were encountered during the purification scale-up phase. We decided to generate bacmids using an improved DH10Bac strain: EmBacY (kindly provided by Prof. I. Berger, University of Bristol, UK). EmBacY strain contains a bacmid that has two baculoviral genes, v-cath and chiA, disrupted. This leads to the reduction of virus dependent proteolytic activity and cell lysis. The v-cath gene encodes for a viral cathepsin-type cysteine protease, V-CATH, which is activated upon cell death by a process that depends on a juxtaposed gene on the viral DNA, chiA, which encodes for a chitinase.^{85,86} Both proteins are involved in the liquefaction of the host insect cells.

The use of this strain for recombinant bacmid generation of human PDE4D3 resulted in an increased yield of stable protein.

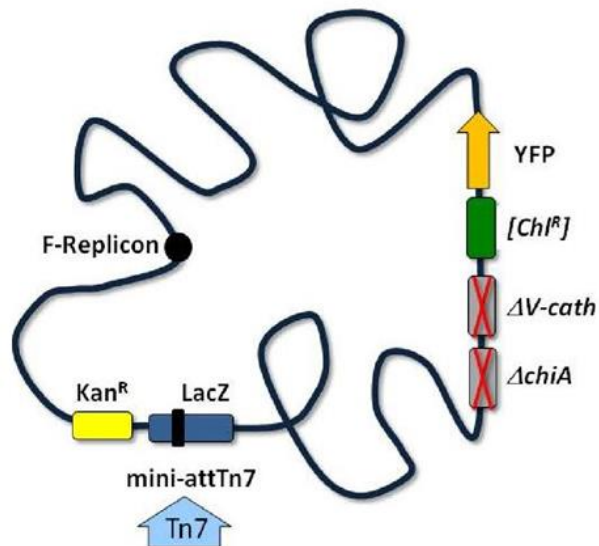


Figure 30: EmBacY baculoviral DNA, adapted from the MultiBac 2011 (MultiBacTurbo) Manual, Alexander Craig & Imre Berger

2.2.6 Purification

The purification of PDE4D3 followed a two steps procedures that includes Ni²⁺ affinity chromatography and ion-exchange chromatography. The protocol was extensively fine-tuned in order to enhance the yield and purity of the sample. In particular, some precautions were found to be of crucial importance:

- **Switching to EmBacY:** as pointed out in the test expression section, the EmBacY strain improved significantly the yield of the target protein (as a ratio versus the degradation products) and its stability over time.
- **Buffer fine tuning:** It is known from the literature that long PDE isoforms are difficult to produce because of their misfolding and degradation-associated instability. In particular, we observed that a minimum concentration of 150 mM NaCl is necessary to avoid precipitation and to improve the stability of the protein during the purification process.
- **Lysate stabilization:** complete inhibitor cocktail (Roche) was introduced to minimize the extensive protease activity that severely affected the very early attempts to purify the protein.

- Ni-NTA improvement: benzonase nuclease was introduced after sonication as a replacement for DNase for DNA degradation. Initially, DNase was utilized for DNA degradation but this turned out to be detrimental as it produced severe column clogging.
- Size exclusion chromatography step removal: This technique was initially used as the second purification step. However, as PDE4D3 does not migrate as a single peak, the size exclusion chromatogram was difficult to analyze and this step resulted in the loss of large amounts of protein.

The final procedure for PDE4D3 purification:

1. Resuspend the cells in lysis buffer at a final concentration of 1×10^9 cells/50 ml buffer (50 mM HEPES pH 7.5, 500 mM NaCl, 10 mM $MgCl_2$, 10% glycerol, TCEP 1 mM). The buffer was supplemented with *Roche complete* inhibitor cocktail (Roche).
2. Lyse the cells by gentle sonication: 3 burst of 3 minutes each at low intensity, with a minute of pause between bursts.
3. Add 2.5 μ L of benzonase nuclease (every 1×10^9 cells) and incubate the lysate for 5 minutes at room temperature
4. Centrifuge at 18000 rpm, for 45 minutes, at 4°C for clarification.
5. Add pre-equilibrated Ni-NTA resin (Qiagen) to the supernatant (700 μ L per 1×10^9 cells). Incubate for 30 minutes at room temperature in agitation.
6. Spin down the resin by centrifuge, 500 rpm for 5 minutes at 4° C.
7. Collect the flow through and reload it on a second Ni-NTA filled column.
8. Wash the resin with 3 column volumes of wash buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 50 mM arginine, 10% glycerol, 1 mM TCEP, 40 mM imidazole)

9. Elute with 4 column volumes of elution buffer (50 mM HEPES pH 7.5, 150 mM NaCl, 50 mM arginine, 10% glycerol, 1 mM TCEP, 300/400 mM imidazole).

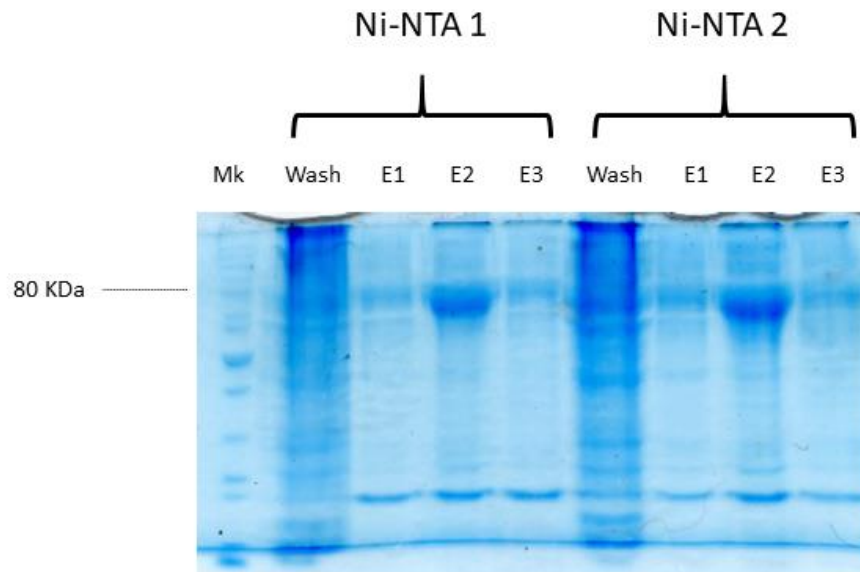


Figure 31: Ni-NTA chromatography of the purification of PDE4D3. Sample loaded on 10% SDS page

10. Overnight dialysis of the elution samples against ION-EX buffer 1 (50 mM HEPES pH 7.5, 150 mM NaCl, 10% glycerol, 1 mM DTT) at 4 °C.
11. Ion-exchange affinity chromatography.
 - a. Elution buffer: (50 mM HEPES pH 7.5, 1M NaCl, 10% glycerol, 1 mM DTT)

ION-EX profiles

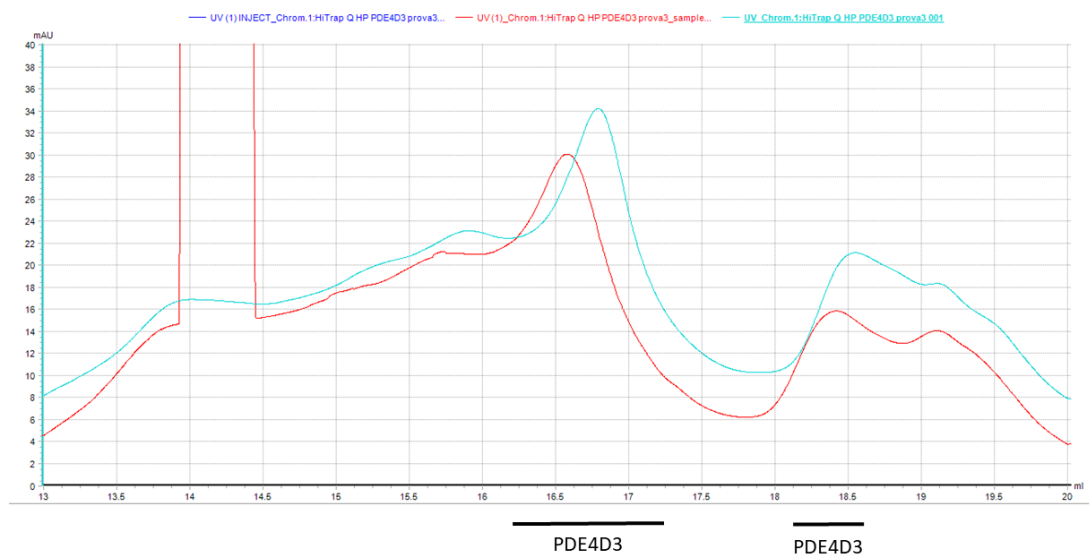


Figure 32: Ion-exchange affinity chromatography elution profiles for PDE4D3 purification.

The protein concentration was then assessed by Bradford assay and brought to 0.025 $\mu\text{g}/\mu\text{L}$ circa.

2.3 PDE4 Catalytic domain X-ray crystallography

Crystallization experiments were carried out using samples of the PDE4D catalytic domain at a concentration of 10 mg/ml (Bradford quantification). All the crystals were obtained by the hanging drop vapor diffusion method, utilizing 24-well plates and 1 ml volumes of the different crystallization solutions in the reservoirs. Crystallization droplets were formed by mixing 1 μl of protein sample with 1 μl of crystallization solution.

2.3.1 Crystal growth - first screening

The design of the crystallization experiments started from a literature analysis of the commonly used reservoirs for PDE4 catalytic domain crystals. Small preliminary crystals were initially obtained after 1 week at 25 $^{\circ}\text{C}$, from conditions containing HEPES pH 7.5 and PEG 3350. A first round of optimization was carried out by changing the concentration of these two components of the crystallization solution, leading to bigger, needle-shaped crystals that, nevertheless, showed no diffraction.

2.3.2 Crystal growth - optimization

Subsequent cycles of crystallization condition optimization were aimed at finding an additive that may allow crystals to grow tridimensionally. Interestingly, the addition of magnesium chloride at a final concentration of 70 to 200 μM in the reservoir changed completely the shape and the

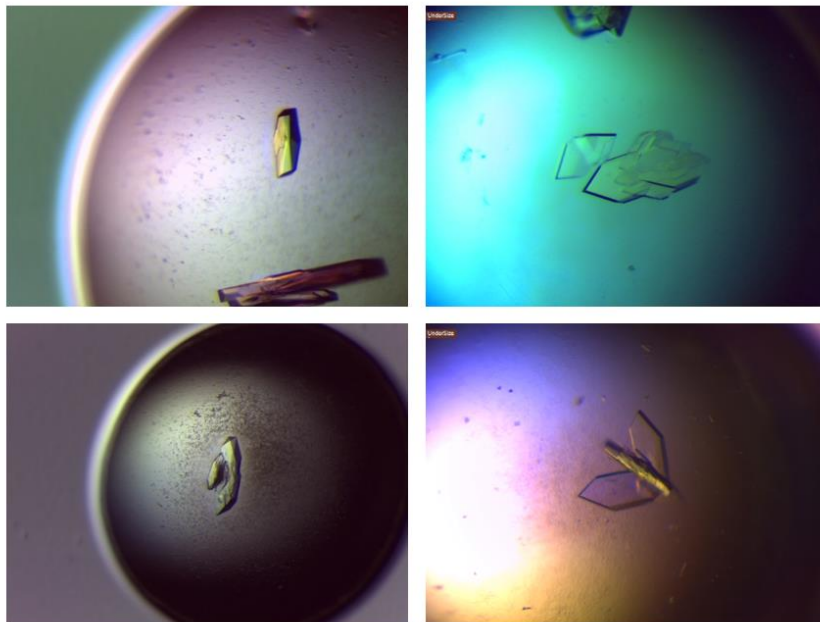


Figure 33: PDE4D catalytic domain crystals

dimension of the crystals, allowing high resolution data diffraction to be collected.

The final crystallization conditions were: 50 mM HEPES pH 7.5, 14-20% PEG 3350, 150 mM MgCl_2 .

The crystallization conditions that yielded crystals of the PDE4 catalytic domain were also tested in co-crystallization experiments with several GEBR compounds. Inhibitors were initially dissolved at a 100 mM concentration in 100% DMSO and subsequently added to the protein sample to reach a final concentration of 0.5-0.6 mM. However, co-crystallization experiment did not result in crystal formation.

2.3.3 Crystal soaking experiments

Owing to co-crystallization failure, soaking experiments were carried out and optimized as an alternative procedure to obtain crystals of the protein-inhibitor complex⁸⁷. Soaking solutions were

prepared by mixing all the components of the reservoir in half their concentrations. The inhibitors were then dissolved in the soaking solutions at their maximum possible concentration (1-6 mM). Soaking experiments were done using fully grown APO crystals of the PDE4 catalytic domain by adding a droplet of the soaking solution to the crystallization droplet. Wells were then sealed again over the former reservoir and plates left on the shelf for 48 hours. Prior to freezing, crystals were periodically monitored over the two days to ascertain whether deterioration occurred. Cryo protectant solutions were prepared with the reservoir condition (half concentration) supplemented with the same concentration of the inhibitor used for the soaking procedure and with 20% ethylene glycol. Crystals were flash frozen in liquid nitrogen and stored for the data collection.

2.3.4 X-ray diffraction collection

X-ray crystallography experiments were done at the Swiss Light Source (SLS) synchrotron, Paul Scherrer Institute (PSI), Villigen, Switzerland. Hardware and experiments details:

- Beamline: X06DA
- Detector: PSI Pilatus 6M
- Wavelength: 1Å
- Data collection temperature: 100K

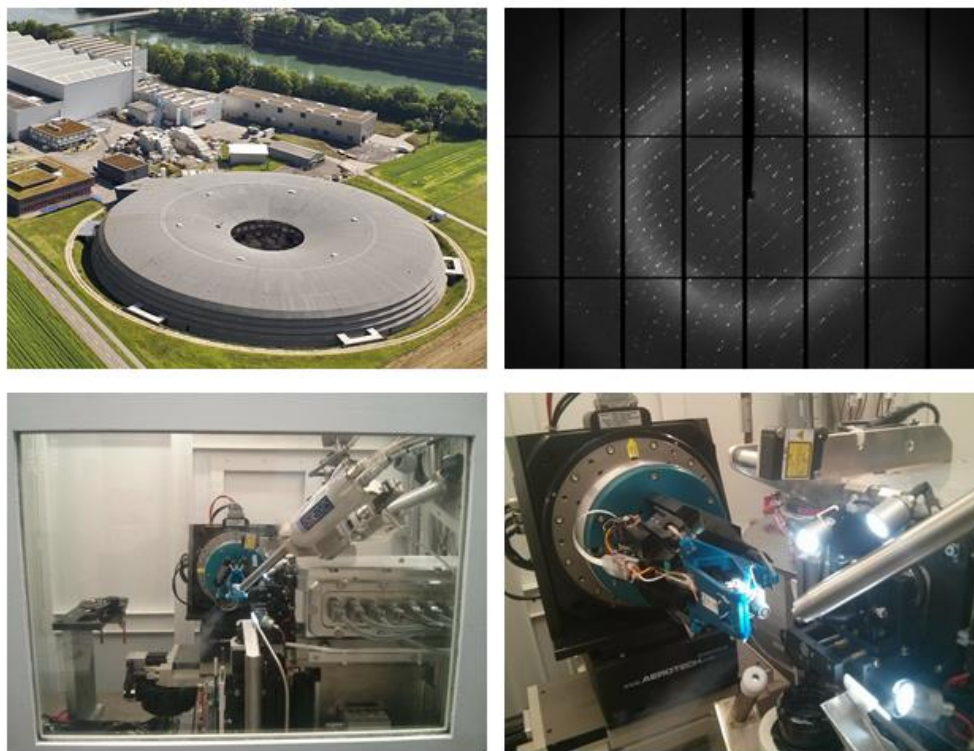


Figure 34: SLS synchrotron ring (top left), PDE4D Cat - GEBR-32a crystal diffraction (top right), X06DA beamline hardware (bottom left and right)

Crystal diffraction was in general collected by rotating the crystal over an ω angle of 180 or 360°, registering images every 0.25 or 0.5°.

2.3.5 Structure solution

In some cases, structure phasing and solution was done starting from the .mtz files already pre-processed using the dedicated software available at the SLS beamline. In some other cases, final .mtz files were re-prepared starting from ASCII.HKL files, using the CCP4 suite software.

Programs utilized for the whole process of structure phasing and solution:

- Import data for scaling: *Combat (CCP4)*.⁸⁸
- Scaling: *Scala (CCP4)*.⁸⁸
- Select reflection for R-free: *Freeflag (CCP4)*.⁸⁸
- Molecular replacement: *Phaser (Phenix 1.9-1692)*.⁸⁹
- Refinements: *Phenix refine (Phenix 1.9-1692)*.⁸⁹

- Ligand building: *Avogadro 1.1.1*.⁹⁰
- Ligand topology generation: *eLBOW (Phenix 1.9-1692)*.⁸⁹
- Model building, completion and validation: *WinCoot 0.8.1*.⁹¹
- Packing inspection: *Chimera 1.8.1, Pymol 1.3*.^{92,93}
- Images recording: *Pymol 1.3*.⁹³

The first dataset was phased using the crystal structure of the PE4D catalytic domain in complex with cAMP (PDB ID: 2PW3). All the subsequent datasets were phased using an optimized monomeric model extracted by our first structure (PDB ID: 6F6U).

The protein structure solution was optimized before adding the compound models in order to obtain the best-phased electron density map. This was done starting by 3 cycles of automated refinement including simulated annealing, and then by adding water molecules, ions and by manually modifying the protein model (each time followed by one cycle of refinement). After every round of model optimization (protein, waters, ions), the F_o-F_c map associated to the ligand was analyzed, the compounds were built using the Avogadro software (generally according to the conformation derived by the map) and manually fitted inside their density before further automated refinement.

For some ligands, several cycles of manual modifications of the compound conformation and refinement were necessary to find the best electron density fit.

2.4 Enzymatic assays

2.4.1 Sensor system

The sensors system⁹⁴ that was used to perform continuous assays of cAMP phosphodiesterase activity follows this coupled scheme:

1. $\text{cAMP} \rightarrow \text{AMP}$ (*PDE4D*)
2. $\text{ATP} + \text{AMP} \rightarrow 2\text{ADP}$ (*Myokinase*)
3. $2\text{PEP} + 2\text{ADP} \rightarrow 2\text{ATP} + 2\text{Pyruvate}$ (*Pyruvate kinase*)
4. $2\text{Pyruvate} + 2\text{NADH} \rightarrow 2\text{NADH}^+ + 2\text{Lactate}$ (*Lactate dehydrogenase*)

The measurements were performed both by absorbance or fluorescence. In the first case, NADH absorbance at 340 nm (molar extinction coefficient = $6220 \text{ M}^{-1}\text{cm}^{-1}$) is followed during the course of the reaction. In the second case, NADH fluorescence decrease at 460 nm (excitation at 355 nm) is followed during the course of the reaction.

2.4.2 Reagents details

- cAMP:
 - Name: adenosine 3',5'-cyclic monophosphate sodium salt
 - Supplier: Sigma Aldrich
 - Stock: 40 μM
 - Final concentration in assay: 4 μM
- PDE4D
 - Name: PDE4D catalytic domain or PDE4D3 (see materials and method, protein expression)
 - Supplier: Produced in lab by *E.coli* heterologous expression (PDE4D catalytic domain) or by *Baculovirus/sf9* heterologous expression (PDE4D3)
 - Stock: 0.0125 – 0.0250 $\mu\text{g}/\mu\text{l}$ circa.
 - Total amount in assay: 0.125 – 0.250 μg . 10 nM circa.
- Myokinase
 - Name: Myokinase, yeast
 - Supplier: Merck
 - Stock: 0.2 U/ μL
 - Final amount in assay: 0.8 U
- ATP
 - Name: Adenosine 5'-triphosphate disodium salt hydrate
 - Supplier: Sigma Aldrich
 - Stock: 3mM
 - Final concentration in assay: 0.3 mM
- PEP:
 - Name: Phospho(enol)pyruvic acid monopotassium salt
 - Supplier: Sigma Aldrich

- Stock: 6 mM
- Final concentration in assay: 0.6 mM
- Pyruvate kinase
 - Name: Pyruvate kinase, rabbit muscle
 - Supplier: Sigma Aldrich
 - Stock: 0.6 U/ μ L
 - Final amount in assay: 0.9 U
- NADH
 - Name: Nicotinamide Adenine Dinucleotide, Reduced Form
 - Supplier: Sigma Aldrich
 - Stock: 1.5 mM
 - Final concentration in assay: 0.15 mM
- Lactate Dehydrogenase
 - Name: L-lactic dehydrogenase, rabbit muscle
 - Supplier: Sigma Aldrich
 - Stock: 0.6 U/ μ L
 - Final amount in assay: 0.9 U
- Buffer
 - For absorbance assays: 100 mM HEPES pH 7.5, MgCl₂ 10 mM, KCl 10 mM, NaCl 100 mM, 1% DMSO.

2.4.3 Experimental Setup

The reactions were set up in transparent polystyrene 96-well plates (Grainer Bio) with a final volume of 250 μ L. Compounds were dissolved in 100% DMSO to a final concentration of 50-100 mM and subsequently diluted to a concentration corresponding to 100x of the final assay condition. The final amount of inhibitor solution pipetted into the well is therefore standardized to 1% of the total reaction volume. The inhibitors were tested in concentrations ranging from 0.01 to 300 μ M, in triplicates. Reaction mixtures (complete of all components except for NADH, cAMP, PDE and the DMSO/inhibitor) were prepared and flash frozen to -80°C in aliquots suitable for IC₅₀ evaluation.

The experiments were performed as follows:

1. Addition of NADH to the reaction mixture
2. Pipette the correct amount of mixture in each well
3. Pipette the correct amount of MilliQ water in each well
4. Addition of 1% DMSO in positive controls.
5. Addition - 2.5 μL of a specific inhibitor in each well
6. Addition - 10 μL of the PDE4 sample
7. Incubation 20' at room temperature
8. Start reaction with addition of 25 μL cAMP
9. Shacking plate for 15 seconds
10. Absorbance/Fluorescence detection (30 minutes) at 37°C

Absorbance or fluorescence were detected and followed using a Tecan Spark 10M plate reader.

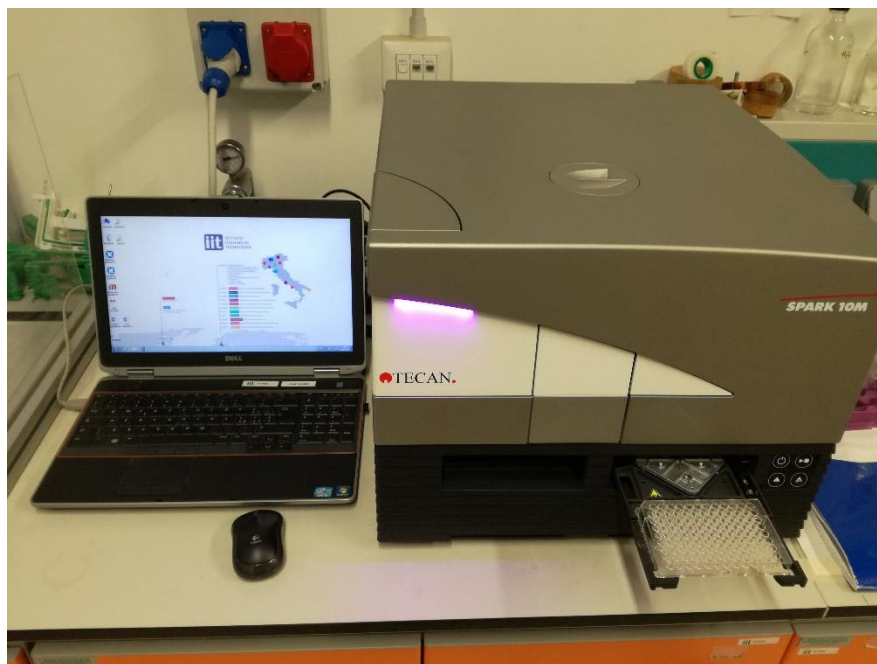


Figure 35: The Tecan Spark 10M plate reader used for the enzymatic assays

2.4.4 Data analysis

Slopes relative to the decrease of absorbance/fluorescence were derived by linear regression of the central, linear part of each curve. Data were processed by SigmaPlot 10.

The activity was calculated as:

$$Activity = \frac{-\Delta Abs/min}{\varepsilon} * V * 10^6 \quad [\mu Mole * min^{-1}]$$

where $\Delta Abs/min$ is the recordered slope, ε is the molar extinction coefficient associated with NADH and V is the volume of the well.

The fractional activity was calculated as:

$$Fractional\ activity = \frac{activity}{positive\ control\ activity}$$

where the *positive control activity* is the average activity of the three positive controls (reactions performed with no inhibitor inside).

Finally, the concentration of compound required to reduce the fractional enzyme activity to half of its initial value (IC_{50}) was determined by plotting the enzyme fractional activity against the logarithm of compound concentration. Curve fitting was performed with a standard dose response curve:

$$y = min + ((max - min) / (1 + 10^{(logIC_{50} - x)}))$$

where y is the fractional activity of the enzyme in the presence of inhibitor at concentration $[I]$, max is the maximum value of y observed at $[I]=0$, and min is the minimum limiting value of y at higher inhibitor concentration.

2.5 Molecular Dynamics simulations

The full length PDE4D structure was modeled using the automatic procedure implemented in the Swiss-model server⁹⁵. The regulatory domains were modeled on the basis of the PDE4B crystal structure (PDB entry: 4WZI)²¹.

2.5.1 Model optimization

Loops were further modeled using scaled molecular dynamics (SMD). Simulations on PDE4D in its holo form were performed with a λ of 0.7 to enhance the sampling of the dihedral angles belonging to the loop region: three simulations of 100 ns each were performed to efficiently flood the conformational space for the loop. This was done by keeping the protein backbone atoms restrained (i.e., using an isotropically applied force constant of 500 kJ mol⁻¹ nm⁻¹) except those of

the loops (i.e., residues 202–253 and 106–136). For each loop, the average starting structure for the subsequent MD simulations was obtained from the cluster analysis performed on all these trajectories. The structure of the full length PDE4D3 in complex with the ligands (GEBR-7b and GEBR-32a) was reconstructed merging the experimental ligand pose (from the X-ray crystal structures) with the model structure of the protein holo form derived as explained above.

The metal coordination was preserved by means of harmonic restraints (with a $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ force constant) applied between the metal ions and the coordinating atoms as determined by the crystallographic structures. Each compound was geometrically optimized prior to MD simulations via a quantum mechanical approach: electron density calculations were performed in Gaussian 09⁹⁶ using the 6-31G* or 6-31g++ basis set at the Hartree–Fock level of theory. Partial charges were derived using the RESP method as implemented in Antechamber, leading via a GAFF parametrization to a complete topological description of each ligand to be used for classical simulations.

2.5.2 Simulations

Protocol details:

- Performed in a GROMACS 4.6.1⁹⁷ implemented in BiKi Life Sciences 1.3 (BiKi Technologies, Genoa, Italy)⁹⁸.
- The force field used for all the classical simulations was Amber ff99SB*-ILDN⁹⁹. All the complexes were first placed in the geometrical center of parallelepiped-shaped boxes with volumes of $\sim 2500 \text{ nm}^3$ and then solvated with ~ 80000 TIP3P water molecules. Some water molecules were replaced with sodium ions to preserve the electro-neutrality of the system according to need.
- The system was minimized with the steepest descent method, followed by equilibration of the restrained protein (isotropic $1000 \text{ kJ mol}^{-1} \text{ nm}^{-1}$ force applied to each heavy atom of the protein backbone) in NPT (up to 400 ps, pressure of 1 atm) and NVT (up to 400 ps) ensembles at 300 K via a standard MD procedure.
- Electrostatics were treated with the cutoff method for short-range interactions and with the particle mesh Ewald method for the long-range interactions (rlist = 1 nm, cutoff distance = 0.9 nm, vdW distance = 0.9 nm, PME order = 4).

- The constant-temperature conditions were provided by using the V-rescale thermostat, a modification from Berendsen's coupling algorithm. The production run was performed for 200 ns for each system.
- All the simulations were set up using the BiKi software package and performed on a set of in-house machines, equipped with two esacore-Intel Xeon processors and two NVIDIA GTX 780 GPUs, for a total of ~800 CPU days.

	GEBR-7b	GEBR-4a	GEBR-20b	GEBR-18b	GEBR-18a	GEBR-32a	GEBR-54	GEBR-26g
PDB ID	6F6U	6F8T	6F8U	6F8V	6F8W	6FDC	6F8R	6F8X
Space group	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21	P 21 21 21
Cell dimensions								
<i>a</i> (Å)	64.83	64.43	64.67	64.56	65.04	64.692	64.67	64.50
<i>b</i> (Å)	98.94	98.57	98.67	98.31	98.72	98.608	98.30	98.55
<i>c</i> (Å)	119.53	118.82	119.64	120.08	120.76	120.165	120.09	119.82
Wavelength (Å)	1	1	1	1	1	1	1	1
Resolution (Å)	1.82	1.8	2.1	1.85	1.6	1.45	1.83	1.95
<i>R</i> _{sym} or <i>R</i> _{merge} (%)	0.106(0.887)	0.045(0.239)	0.068(0.285)	0.096(0.851)	0.073(1.033)	0.06(0.837)	0.102(0.784)	0.124(0.604)
<i>I</i> / σ <i>I</i>	10 (1.7)	22.75 (5.8)	16.5 (5.2)	11.2 (2.2)	12.3 (1.7)	21.5 (3.4)	9.1 (2.0)	14.2 (3.9)
<i>CC</i> _{1/2}	0.99 (0.77)	0.99 (0.97)	0.99 (0.94)	0.99 (0.85)	0.99 (0.68)	1.00 (0.87)	0.99 (0.69)	0.99 (0.92)
Completeness (%)	98.69(92.86)	99.89(98.95)	99.8 (99.4)	99.4 (97.5)	100 (99.9)	100 (100)	99.1 (96.7)	100 (100)
Multiplicity	6.5 (6.1)	6.6 (6.6)	4.4 (4.3)	6.5 (6.2)	6.6 (6.4)	13.2 (12.4)	5.1 (4.6)	13.0 (11.4)
No. of reflections	67790	71439	45081	64377	196940	263295	67722	107507
<i>R</i> _{work} / <i>R</i> _{free}	0.18/0.23	0.17/0.21	0.17/0.22	0.20/0.25	0.19/0.22	0.18/0.20	0.18/0.23	0.16/0.20
No. of atoms:								
Protein	5298	5334	5296	5295	5378	5345	5269	5281
Inhibitor	56	58	60	62	58	58	58	58
Water	581	648	510	554	522	504	636	502
Ions	5	7	7	5	7	7	4	7
Other	12	0	0	0	0	8	0	12
Average B-factors (Å ²):								
Protein	26.07	21.98	24.83	23.13	27.97	22.29	22.83	23.35
Inhibitor	28.77	22.83	33.23	24.76	24.32	26.79	30.58	24.75
Water	34.30	32.20	32.50	32.00	36.23	31.16	32.11	33.03
r.m.s.d								
Bond lengths (Å)	0.006	0.007	0.007	0.008	0.006	0.006	0.007	0.007
Bond angles (°)	0.9	0.98	1.0	1.0	1.0	1.0	1.0	0.9
Ramachandran								
Most favoured (%)	97.02	96.22	96.98	97.18	96.69	96.68	98.62	98.17
Additional all. (%)	2.66	2.83	2.54	2.19	2.37	2.53	1.38	1.83
Disallowed (%)	0.31	0.94	0.48	0.63	0.95	0.69	0	0

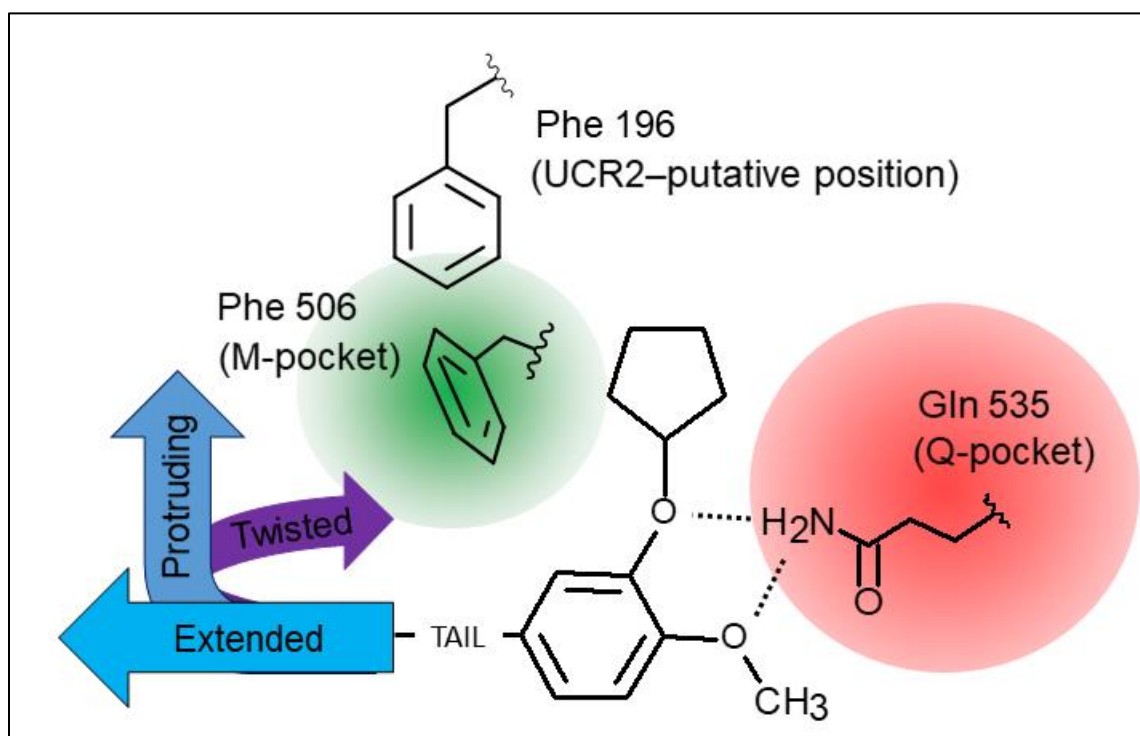
Table 2: crystallographic data

As described more extensively in the introduction, within the PDE catalytic domain, the active site consists of a 15 Å-deep pocket that is lined by highly conserved residues across the different PDE isoforms. The catalytic pocket can be subdivided into three contiguous sectors, commonly referred to as the metal binding pocket (M), the Q switch and P clamp pocket (Q) and the solvent-filled side pocket (S)⁴⁷ (see Figure 11). The metal binding pocket nests two metal ions (Zn and Mg), which are octahedrally coordinated by conserved His and Asp residues as well as water molecules. Next to the M pocket is the Q switch and P clamp pocket, featuring a Gln residue that acts as the purine-specific coordination site that guides the docking of cAMP into the catalytic site. Moreover, the Q pocket features two conserved hydrophobic residues (Ile502 and Phe538) that stabilize the substrate (P clamp). Finally, the S pocket is filled with an extended network of water molecules. To date, all known catechol-based inhibitors adopt a similar binding mode whereby the aromatic ring of the catecholic portion of the ligands is stabilized by the hydrophobic clamp formed by Phe538 and Ile502, the former engaging in a π - π interaction with the aromatic ring of the catechol⁴⁷ while the latter forming extended van der Waals contacts with it. Most if not all of the previously characterized catechol-based ligands make stabilizing contacts with portions of the M and the Q

pockets, while only recently catechol-based inhibitors that extend into the S-pocket have been described (PDB ID: 5OHJ).¹⁰⁰

In the structures described herein, the binding conformation of the catecholic portion is highly conserved and virtually identical to that observed in the crystal structure of the Rolipram complex (PDB ID: 1OYN). However, notable structural differences can be observed when comparing the tail conformation of each ligand. Sorting our crystal structures according to the direction of the tails of the ligands, we can identify three main conformational ensembles: protruding (GEBR-7b, GEBR-4a, GEBR-32a, GEBR-54, GEBR-20b), twisted (GEBR-18a, GEBR-18b) and extended (GEBR-26g). Although all the crystal structures show two independent catalytic domain-ligand complexes in the asymmetric unit, in most cases little if any structural differences could be observed in the conformation of the same ligand in the two independent complexes, thus suggesting a relatively stable and conserved binding conformation for each inhibitor.

Figure 37: Schematic representation of the conformationally-different families of GEBR library compounds



The first ensemble (protruding) can be split into two chemically different groups, where group 1 is represented by GEBR-7b-related compounds while group 2 consists of those GEBR-54-related molecules that are characterized by a pyrazole linker carrying a 2-hydroxypropyl chain in N1 (see

compound table in Figure 36). Within all the protruding ligands, the tails develop towards the external part of the active site, extending into the S pocket where they make water-mediated contacts with the lining residues. In particular, all these ligands contact Phe506 through both their cyclopentane moiety and their morpholine portion. Further stabilizing interactions are provided by water-mediated contacts inside the M pocket or at the interface between the M and the S pocket.

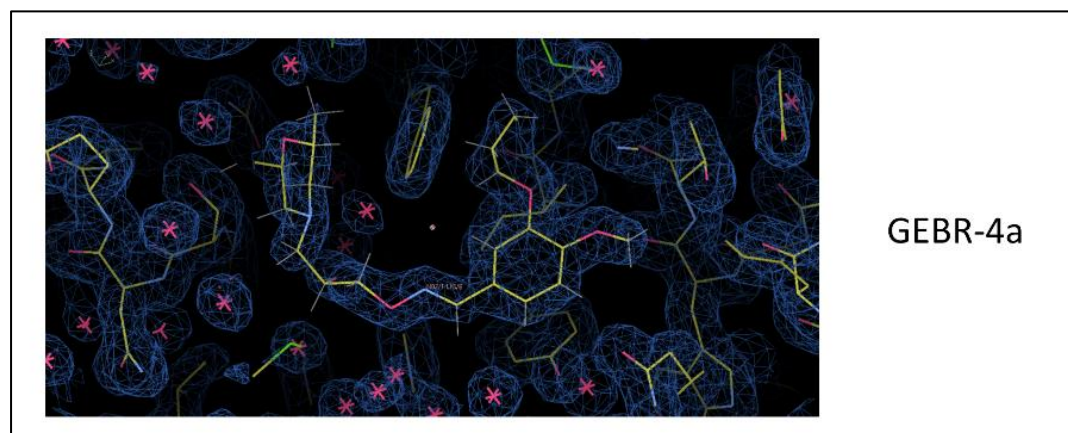


Figure 38: Detail of the structure of the PDE4D catalytic domain - GEBR-4a complex. Map countoured at the 1.0 σ -level.

Interestingly, the high resolution structure of GEBR-54 shown herein displays a conformation that contradicts the previously hypothesized pose of the ligand inside the catalytic pocket obtained by molecular dynamics on the catalytic domain⁸⁰. In particular, a dramatic difference can be appreciated in the direction of the tail as the morpholine moiety fits into the S pocket, almost 8 Å away from Met439, its predicted interactor. It is worth mentioning that GEBR-54 adopts two different conformations in the two independent monomers: one common to all the protruding ligands and developing into the S pocket, the other featuring close contacts between the morpholine moiety and a hydrophobic area formed by Phe538 and Ile542. In the second conformation the tail of the ligand is shifted towards the entrance of the catalytic pocket, suggesting an even higher degree of flexibility.

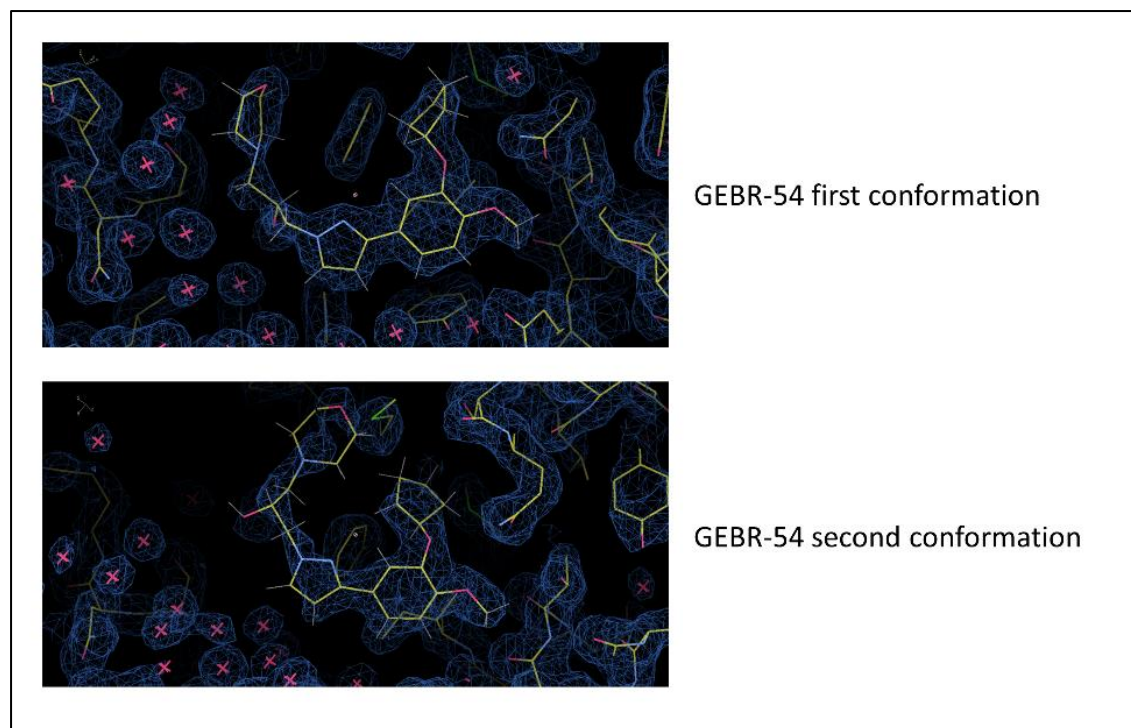


Figure 39: Details of the structure of the PDE4D catalytic domain - GEBR-54 complex. Double conformation of the ligand – map contoured at the 1.0 σ -level.

The second ensemble (twisted) is represented by those GEBR-54-related ligands that are characterized by a pyrazole linker carrying a propionyl chain on the N1, thus likely featuring a more conformationally-constrained tail (GEBR-18b and GEBR-18a). Despite the availability of the S pocket, the twisted ligands show a sharp turn in their structure, resulting in the formation of stabilizing intramolecular interactions with the pyrazole ring. Indeed, the morpholine and the pyrazole moieties, which are connected through a three-atom linker, are nearly parallel to each other. Moreover, morpholine makes intermolecular hydrogen bonding contacts with water molecules in the M pocket as well as hydrophobic interactions with Phe506 and Ile502. These hydrophobic contacts are further enhanced when dimethyl morpholine is present (GEBR-18b). In contrast with the protruding ensemble, here the tails are found in the M pocket over the first solvation sphere of the catalytic ions. In GEBR-18b and GEBR-18a the nitrogen atom of the pyrazole points toward the M pocket, forming a water-mediated interaction with residues Tyr325 and Asp482. Interestingly, this hub water molecule appears to be conserved in all crystal structures including the cAMP-PDE4 complex (PDB code: 2PW3).¹⁰¹ The carbonyl in the linker portion is engaged in a hydrogen bonding network involving two water molecules that interact both with Asp396. One of these water molecules also belongs to the first solvation shell of Mg^{2+} .

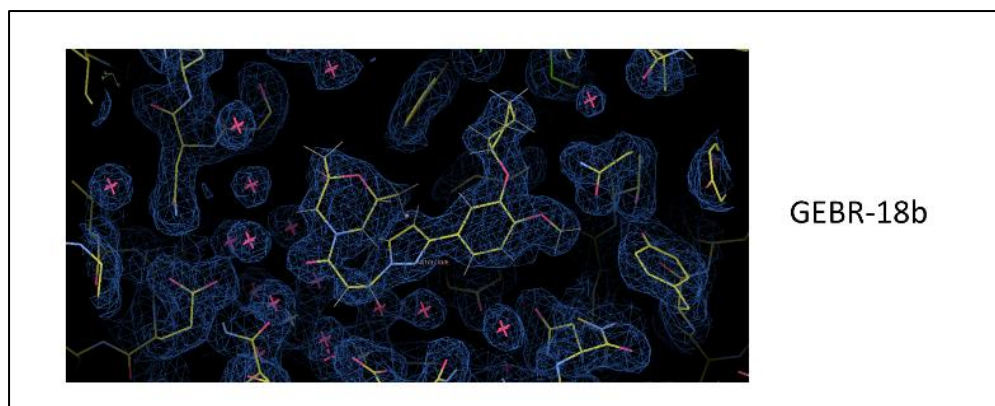


Figure 40: Detail of the structure of the PDE4D catalytic domain – GEBR complex-18b. Map contoured at the 1.0 σ -level.

GEBR-26g is the only member of the third conformational ensemble (extended). Its flexible isoxazoline-containing tail extends at the interface of the M and the S pockets, untwisted and flattened into the active site. Interestingly, the heteroatoms (O, N) of the central isoxazoline ring point towards the external part of the pocket and are not involved in any stabilizing interactions. This allows the tail to lay over the M pocket, with its two branching terminal moieties forming hydrogen bonding contacts with the backbone amide of Asn375 on one side and water mediated contacts with Glu505 and His370 on the other side. The carbonyl moiety in the linker region forms a water-mediated contact with both Met439 and Glu396. It is important to note that the crystal structure is consistent with only one enantiomer of the GEBR-26g racemic mixture docked inside the pocket. In fact, the electron density nearby the chiral carbon of the isoxazoline ring shows a directionality that is incompatible with the other enantiomer.

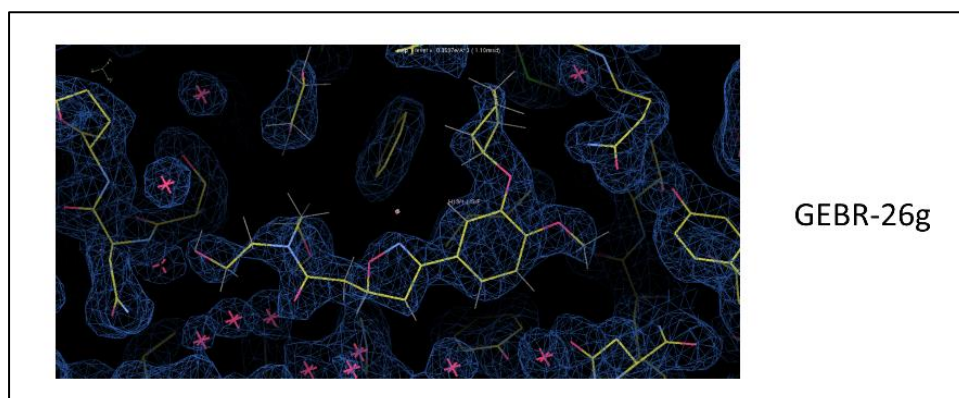


Figure 41: Detail of the structure of the PDE4D catalytic domain - GEBR-26g complex. Map contoured at the 1.0 σ -level.

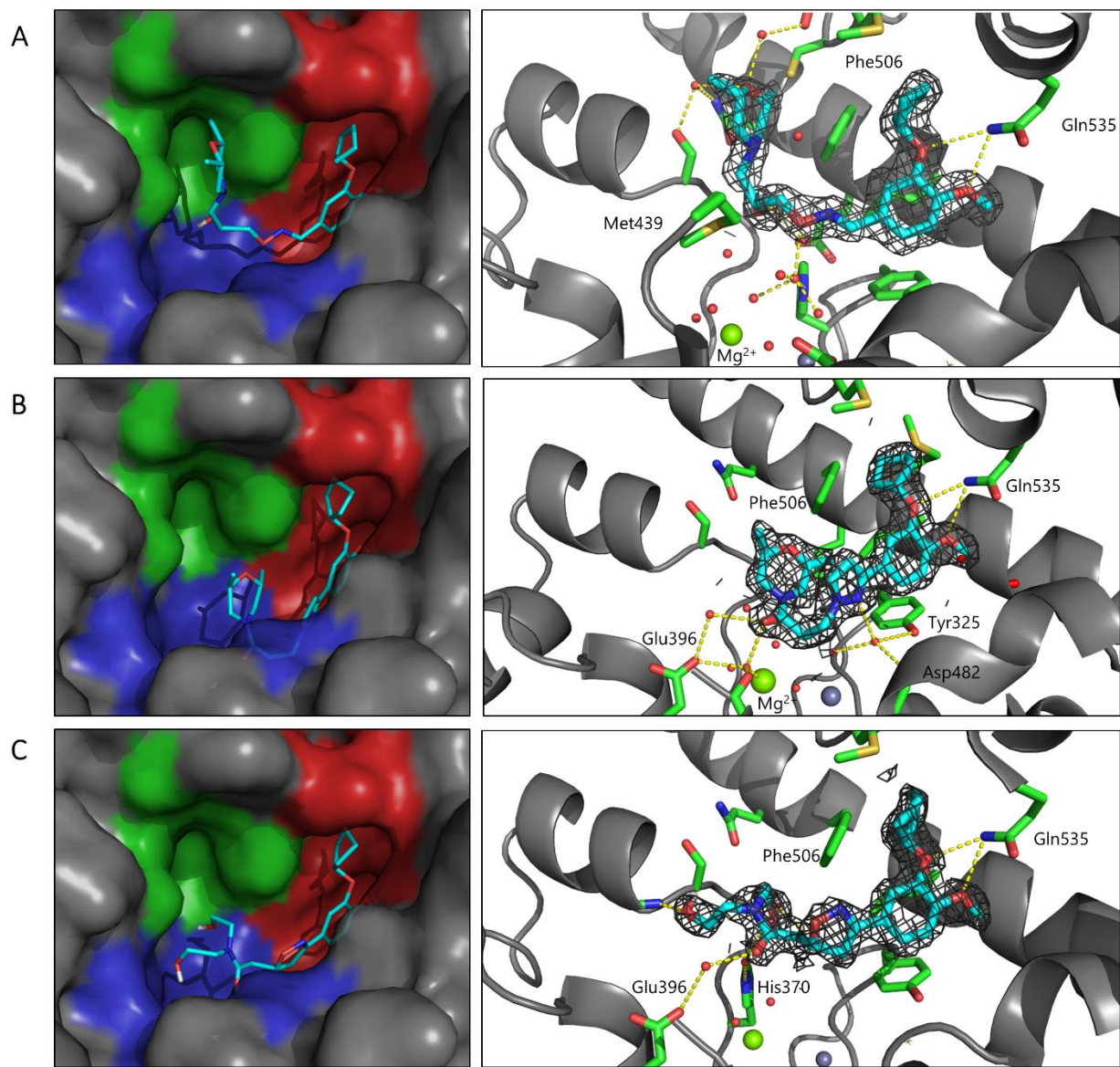


Figure 42 Binding conformations of three prototypical members of the GEBR library: (A) GEBR-4a, protruding, (B) GEBR-18b, twisted, and (C) GEBR-26g, extended.

Within all the conformational ensembles described herein, the protruding ligands appear to be of particular interest due to the fact that they extend into the S pocket and they develop towards the entrance of the active site. Interestingly, GEBR-7b, GEBR-54 and GEBR-32a, which have been shown to have the most interesting pharmacological profile^{72,80,82} belong to this subset of protruding ligands. Potentially, this feature may allow them to interact with the portion of the

regulatory domain that caps the catalytic pocket in the intact protein, thus possibly contributing to their inhibitory mechanism. Therefore, to investigate this further, we carried out inhibition assays on both the full length PDE4D3 protein and the catalytic domain alone in order to assess any differential inhibitory behavior for all the studied ligands.

Comp.	IC ₅₀ (PDE4D Cat)	IC ₅₀ (PDE4D3)
GEBR-7b	16 ± 2.2	11 ± 2
GEBR-20b	0.8 ± 0.2	0.6 ± 0.2
GEBR-4a	7 ± 2	2.1 ± 0.3
GEBR-11b	0.9 ± 0.1	0.4 ± 0.1
GEBR-26g	17 ± 5	3.5 ± 0.8
GEBR-54	20 ± 3	4.6 ± 1
GEBR-32a	2.3 ± 0.7	1 ± 0.2
GEBR-18b	16 ± 5	4.8 ± 0.7
GEBR-18a	23 ± 4	5 ± 0.9

Figure 43: IC₅₀ (μM) Values Measured against the catalytic domain only (Cat) and against the full length isoform (PDE4D3)

Interestingly, all the ligands described herein are endowed with higher IC₅₀ values than that reported in the literature for Rolipram, both towards the full PDE4D3 isoform and towards the catalytic domain alone. Moreover, unlike Rolipram, they do not show full inhibition in the range of concentrations tested, suggesting that despite the presence of the catecholic portion and the conserved binding mode,

the steric hindrance provided by the long tails may impact on the ligand binding efficiency. It is worth stressing that, as pointed out by Burgin et al.³¹, full inhibition is not a strict requirement as partial inhibition may reduce target-based toxicity while maintaining basal levels of cAMP signaling.

The complex PDE4 activation mechanism involves phosphorylation at several different sites. It has been demonstrated that the phosphorylation of Ser54 alters interactions of the UCR1 and UCR2 domains and causes UCR2 to adopt the open active conformation while the phosphorylation of Ser579 stabilizes the closed conformation. Therefore, we performed inhibition assays on a PDE4D3 construct carrying the phospho-mimicking mutations Ser54Glu and the phospho-abolishing mutation Ser579Ala to produce a constitutively active form of the enzyme (see the materials and methods section).

IC₅₀ values against the PDE4D3 isoform and against the catalytic domain only are shown in Figure 44. In all cases, there is a > 1.5 fold increase in IC₅₀ when the compounds are tested on the full length system, the only exception being GEBR-7b and its fluorinated analogue GEBR-20b, for which a lower than 1.5-fold increase in IC₅₀ is observed. Interestingly, in the context of type 1 scaffold (Figure 36) these are the compounds that feature a shorter and less flexible tail. For all the

other compounds, the fold increase in potency in going from the catalytic domain to the PDE4D3 isoform ranges from 2.3 to 4.8. GEBR-26g, which displays an extended conformation at the interface between the M and the S pocket and therefore does not protrude towards the outside of the catalytic pocket, shows a 4.8-fold increase in potency when tested against the PDE4D3 isoform. It therefore appears that, within our subset of GEBR compounds, no clear correlation can be drawn between the binding conformation of the ligand in the pocket and its increase in potency on the full length protein, although longer compounds appear to achieve a better inhibition of the full length protein.

Bruno *et al.*²⁶ showed that GEBR-4a is able to inhibit the long PDE4D3 isoform three times more than the short isoform PDE4D2 (IC_{50} 2.9 vs 9.2 μ M). Consistently, we observed that the same compound inhibits PDE4D3 three times more than the catalytic domain alone (IC_{50} 2.1 ± 0.3 vs. 7 ± 2 μ M). Likewise, GEBR-7b has been shown by Bruno *et al.* to inhibit both PDE4D2 and PDE4D3 similarly²⁶. In our tests, GEBR-7b as well as its fluorinated counterpart GEBR-20b also shows no significant differential inhibition. This may suggest that the extent by which the tail of the ligands protrude does impact on the ability to provide an interaction with the trans-capping portion of the dimeric enzyme, the compounds with longer protruding tails being able to provide a better chemical environment for the capping helix.

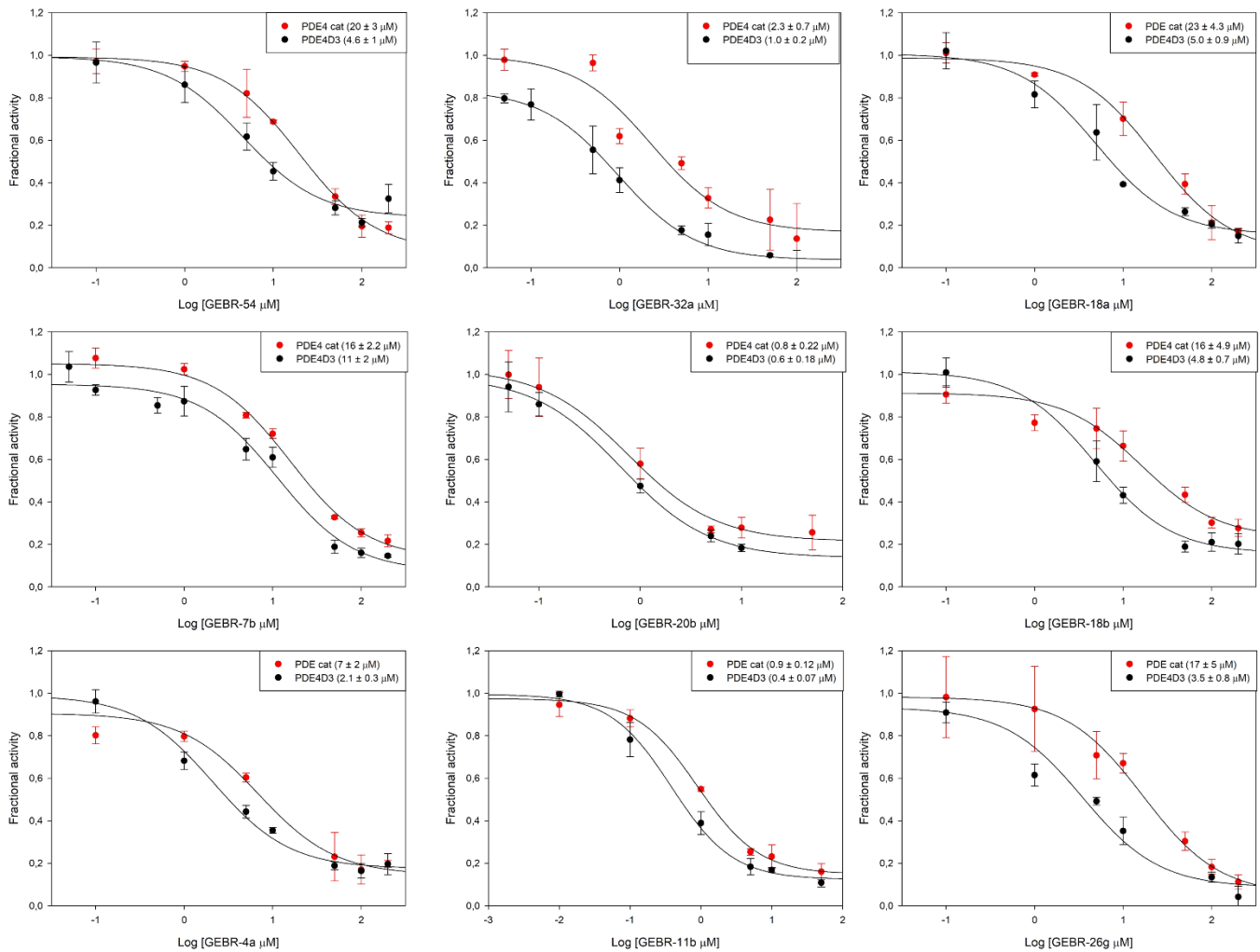


Figure 44: IC_{50} curves for the different compounds relative to the PDE4D3 isoform (black dots) and the PDE4 catalytic domain only (red dots). The reported data are the mean values of three replicates \pm SD

3.1 Inside the GEBR7b-like compounds: role of the longer chain and of the hydroxyl moiety

In the context of GEBR-7b scaffold, we have evaluated the effect of a longer and more flexible ligand. Indeed, as the UCR2 domain is considered to be flexible and subject to conformational variations, a longer and more adaptable ligand tail could fit and better adapt to the target. This

analysis is based on the crystal structure of the GEBR-7b and GEBR-4a complexes, the latter featuring one carbon atom-longer chain connecting the morpholine ring to the catecholic phenyl and a hydroxyl group instead of the carbonyl.

As expected, while the catechol portion of the two compounds are perfectly superimposable in the two crystal structures, the tails protrude from the catalytic pocket in a slightly different conformation. In fact, GEBR-4a appears to better fit the pocket: the portion between the nitrogen of the linker region and the carbon carrying the hydroxyl group is nearly 2 Å lower into the pocket than the one of GEBR-7b. This lowering inside the catalytic pocket results in an augmented set of interactions of the ligand with the local polar chemical environment. This picture is also supported

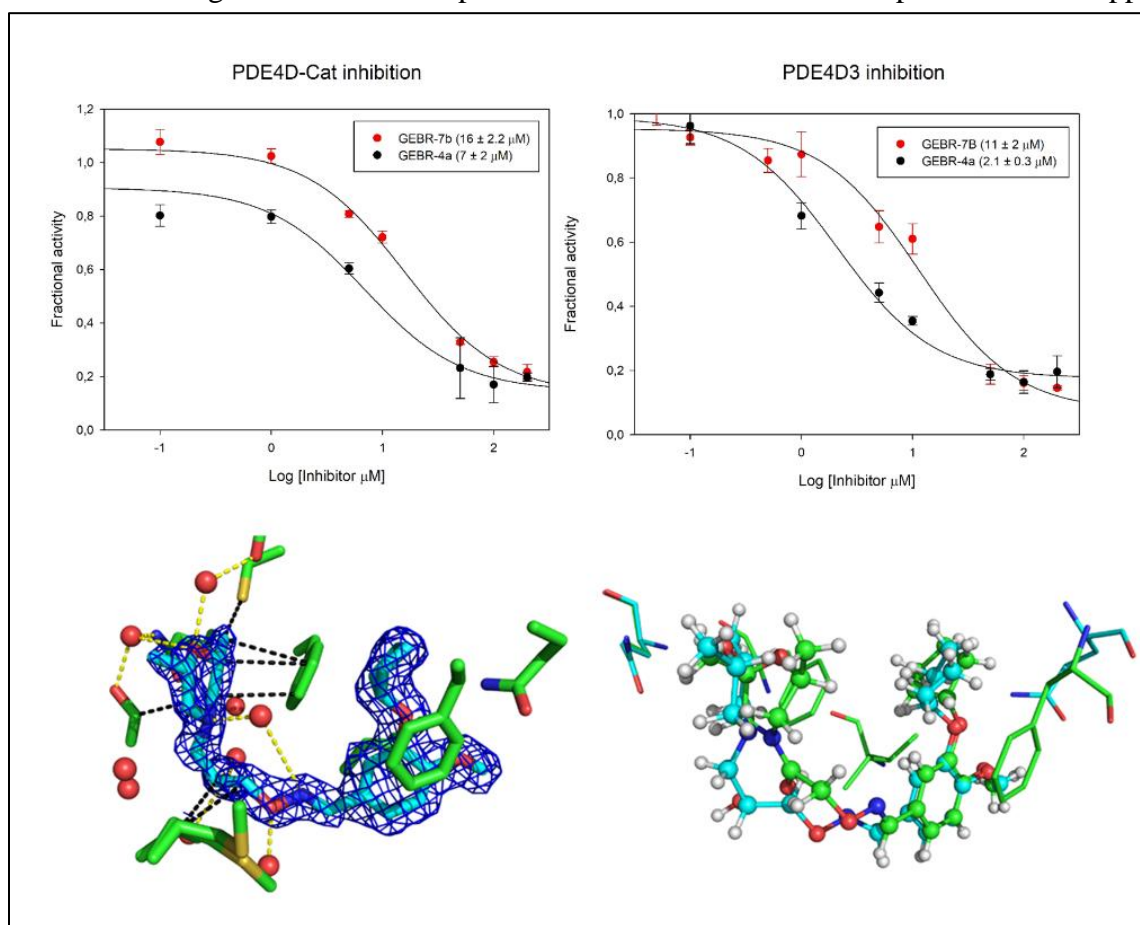


Figure 45:

Top panel - Comparison between the inhibition by GEBR-7b and GEBR-4a of the catalytic domain (left) and of the PDE4D3 isoform (right).

Bottom right panel – Details of the superimposition of the crystal structures of the PDE4 catalytic domain-GEBR-7b (green model) and PDE4 catalytic domain-GEBR-4a (cyan model) complexes.

Bottom left panel – Crystal structure of the PDE4 catalytic domain-GEBR-4a complex with details of the augmented set of interactions

by the inhibition assays that were done against the PDE4D catalytic domain for these two compounds. Indeed, GEBR-7b shows an IC_{50} of $16 \pm 2 \mu\text{M}$ while GEBR-4a has an IC_{50} of $7 \pm 2 \mu\text{M}$. This effect is also more pronounced in PDE4D3 where GEBR-7b shows an IC_{50} of $11 \pm 2 \mu\text{M}$ while it is $2 \pm 0.3 \mu\text{M}$ for GEBR-4a. As previously pointed out, with the longer and more flexible chain the appearance of a differential inhibition profile between the PDE4D catalytic domain and PDE4D3 can be appreciated. While GEBR-7b provides virtually the same level of inhibition to the two forms of the enzyme, GEBR-4a shows a 3-fold increase in the inhibition of the full length PDE4D3 compared to the PDE4D catalytic domain. The whole picture suggests that a longer and more flexible chain may efficiently adapt to the chemical environment of the active site, in particular when a regulatory domain of the enzyme is capping the catalytic pocket.

3.2 Role of the fluorine atoms

The library described herein accounts for several compounds that differ for just the presence of a CH_3 or a CHF_2 group in the catecholic portion. The CHF_2 group is present also in the PDE4

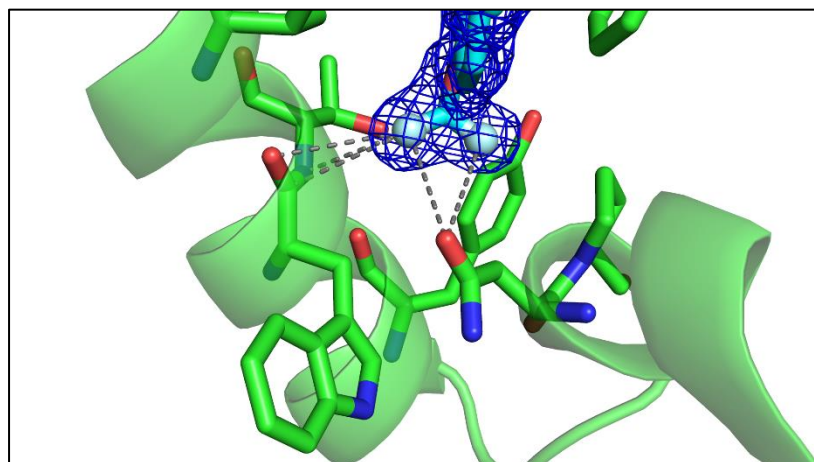


Figure 46: Multipolar interactions in the Q pocket involving the difluoromethoxy moiety of the ligands and Trp498 and Ans487. The $2F_o-F_c$ map is shown at the 1.0σ level

approved inhibitor roflumilast, which is utilized for the treatment of chronic obstructive pulmonary disease (COPD).

In our compounds, the net effect of fluorination appears to be significant. Of the three fluorinated/non-fluorinated pairs that we screened (GEBR-

7b/20b, GEBR-4a/11b, GEBR-54/GEBR-32a), the fluorinated inhibitors always show a lower IC_{50} value than the non-fluorinated ones, both when tested against the PDE4 catalytic domain alone and against the full length PDE4D3 isoform. As previously reported for the PDE4-roflumilast complex structure⁴⁷, also in the GEBR compounds the fluorine atoms make hydrophobic interactions with the protein inside the Q1 pocket. Interestingly, while one of the fluorine atoms interacts with the carbonyl oxygen of Trp498

(3.4-3.5 Å), the two of them interact simultaneously with the side chain oxygen of Asn487 (3.6 Å), forming a multipolar contact network that is commonly observed when fluorinated compounds interact with proteins.¹⁰²

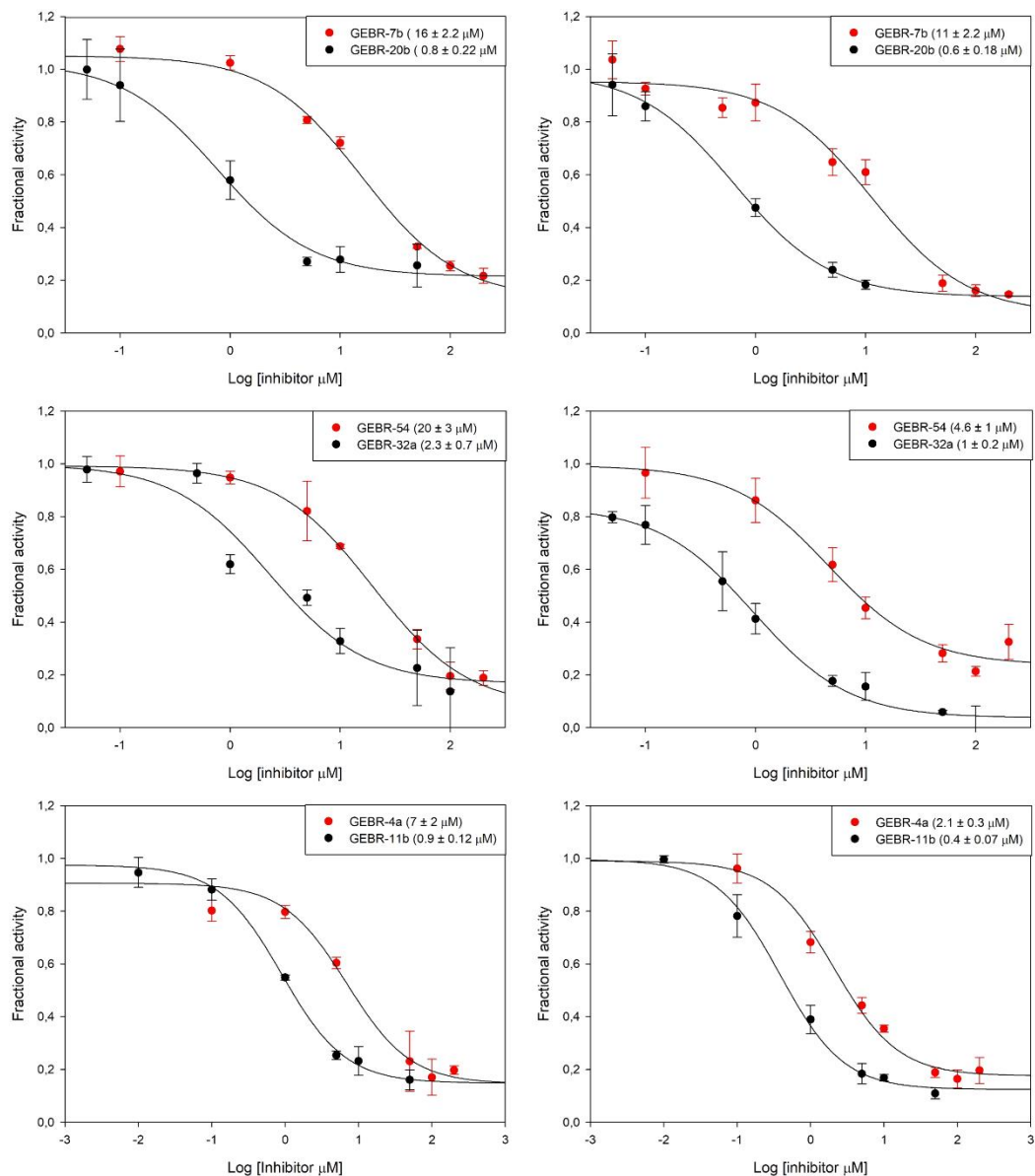


Figure 47: Comparison of the IC₅₀ curves of the three pairs of fluorinated (black dots) and non-fluorinated (red dots) compounds. IC₅₀ values are measured relatively to the PDE4D catalytic domain (left column) and to the PDE4D3 isoform (right column).

3.3 GEBR-32a tautomerism

In the GEBR-32a complex structure, one of the two monomers features the expected ligand in the protruding conformation, while the second one displays a tautomeric form in which the tail of the ligand is attached to the N2 of the pyrazole ring. This tautomeric form, which also adopts a protruding conformation, is a secondary product of the synthetic process. In fact, as already reported in the literature¹⁰³, *N*-substitutions on the pyrazole ring often provide a mixture of two regioisomers. Whereas their distribution is generally unpredictable, the N1-substituted derivative is often the major one, as in our case. The presence of two tautomers, not highlighted by preliminary investigations but clearly observed in the electron density map of the crystal structure reported here, has been confirmed by a more accurate NMR analysis, which also allowed the identification of the mixture composition. NMR data provided evidence that also the oxirane derivative **1**, which is the GEBR-32a precursor, was obtained as an inseparable mixture 1a/1b, thus resulting in GEBR-32a being also obtained as a mixture GEBR-32a/GEBR-32a^{taut} (5:1) (Figure 48). The structure of **1a** was assigned to the major oxirane tautomer by comparing the more intense peaks in the ¹H NMR spectrum of the analogue 3-(3-cyclopentyloxy-4-methoxy-phenyl)-1-oxiranylmethyl-1*H*-pyrazole already reported⁸⁰ and those of the analyzed mixture. Consequently, the isomeric structure of 1b was assigned to the minor tautomer. The ratio 1a/1b (2.1:1) was calculated by comparing the integral values of the peaks of H_A of CH₂O-epoxy group at 2.53 ppm and 2.58 ppm respectively. In the case of the GERB-32a mixture, the GEBR-32a structure was assigned to the major tautomer by comparing the chemical shift of the more intense peaks with those of the precursor 1a. The ratio (5:1) was estimated by comparing the integral values of the peaks of the CH-O group of the cyclopentyl substituent at 4.91 ppm and 4.83 ppm of the major and the minor tautomer, respectively.

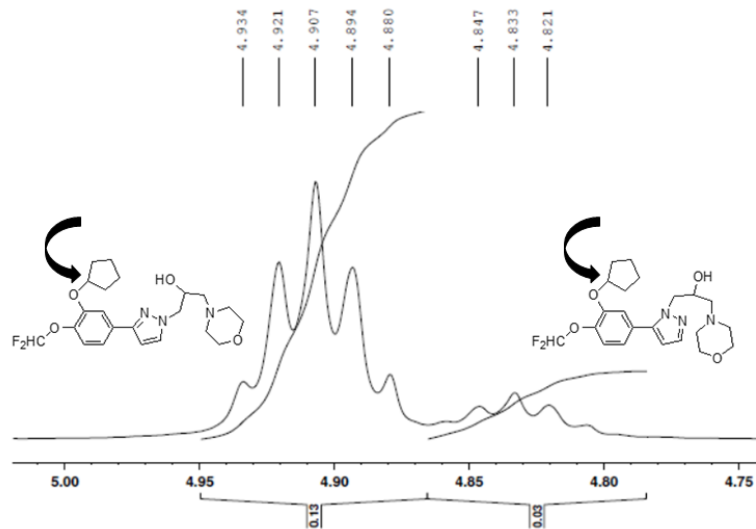


Figure 48: The NMR signals used to determine the GEBR-32a/GEBR-32a-tautomer ratio.

3.4 Molecular dynamics

To gain an insight into the dynamic behavior of the ligands described herein in the context of a long PDE4D isoform that includes the regulatory domains UCR2 and part of UCR1, 200 ns molecular dynamics (MD) simulations of the complexes with GEBR-7b and GEBR-32a were carried out. The choice of these compounds has been guided by their very promising pharmacological profiles observed in animal models. Furthermore, they feature two different scaffolds and they are both protruding. Overall, the pose provided by the crystallographic structures is maintained in both cases during the entire course of the simulations. The relatively low ($< 3\text{\AA}$) RMSD of the binding sites of the ligands, calculated for all their atoms, reveals that the local structure of PDE4D involved in the harboring of the ligand is retained during the simulation time.

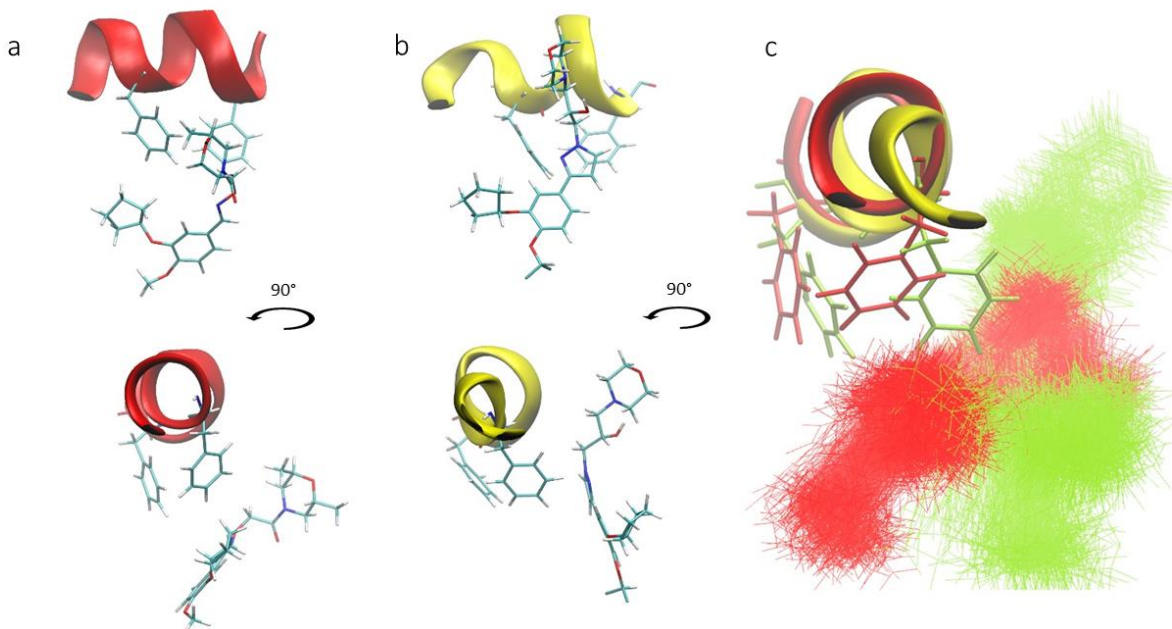


Figure 49: Average orientation of (a) GEBR-7b and (b) GEBR-32a in the binding site with respect to the UCR2 domain at the end of MD simulations. The helix moiety of UCR2 is represented as ribbons (red for GEBR-7b and yellow for GEBR-32a) spanning the residues 98-10

The RMSD of the ligands, calculated on the backbone atoms of the binding site residues only, demonstrate the low positional variability of both ligands in the catalytic pocket. In the PDE4D dimer, while GEBR-32a displays the same conformation for both the monomers together with a low RMSD for the catalytic pockets, GEBR-7b assumes two different final conformations. One (conformation A) closely resembles the experimental structure in terms of both the tail orientation and the distance to the neighboring residues, whereas the other (conformation B, not shown) is markedly different to the experimental one. Notably, a pose similar to conformation B, had been previously hypothesized for GEBR-54 in a recent publication⁸⁰ reporting a putative docking-based binding mode of the ligand in the catalytic domain of PDE4D. It is conceivable that GEBR-7b may be able to explore conformations that are less energetically favorable and hence experimentally inaccessible or poorly accessible due to kinetic reasons.

By visual inspection, the comparison between the binding mode of GEBR-7b and GEBR-32a reveals for the latter an average final position closer to the UCR2 capping helix, with the compound assuming a concave conformation under Phe196, whose implication in PDE4D drug development has been previously described by Burgin and coworkers.³¹ A bundle of frames reported from the

simulation shows the calculated conformational differences between GEBR-7b and GEBR-32a. In both cases, the conformation of the ligand tail is stabilized by the orientation of Phe196 which, together with Phe506 and Phe538 (Figure 50) provides a harboring hydrophobic groove.

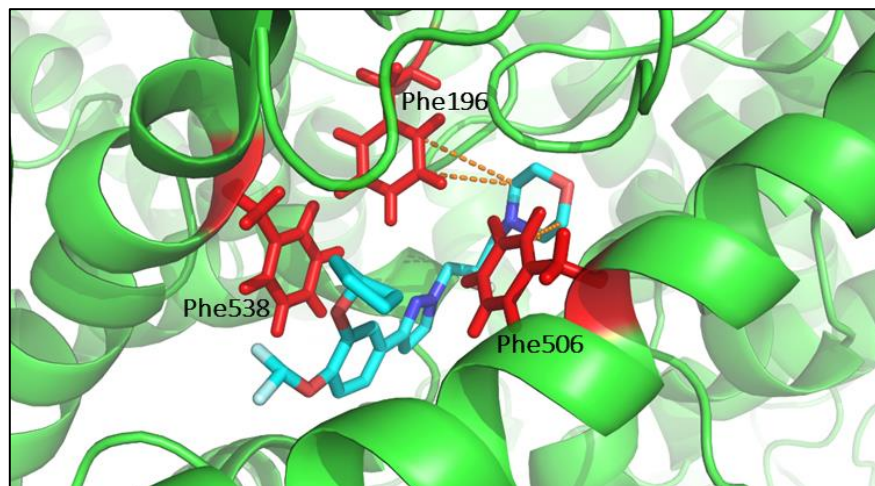


Figure 50: Theoretical model (frame taken from GEBR-32a simulation) of the hydrophobic groove that is formed at the interface between the catalytic domain and UCR2.

To the best of our knowledge, this is the first MD simulation of a long PDE4D isoform that complements experimentally derived data. To date, owing to the difficulty in obtaining experimental structural data for the full-length form of the protein, MD may be

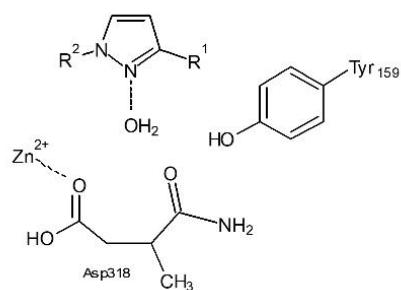
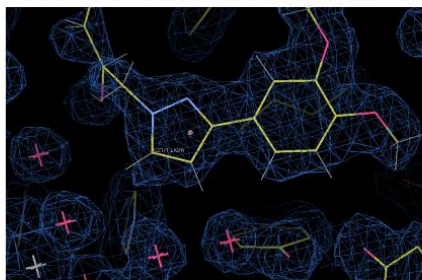
considered, the best *in silico* technique that can be used to validate the stability of the crystallographic conformation in the presence of the capping regulatory domain UCR2.

3.5 Pyrrole compounds

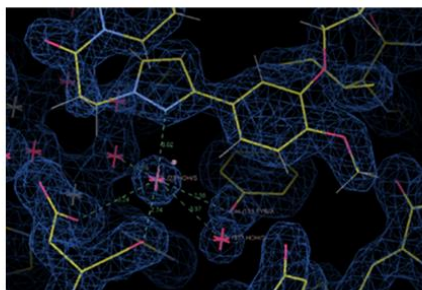
A new class of GEBR compounds has been synthesized during this project. This subset features a GEBR-54-like scaffold in which the pyrazole moiety has been substituted with a pyrrole one. Preliminary data (unpublished) suggest that this substitution does not affect the conformational behavior of the inhibitors and protruding and twisted compounds can be developed by changing the linker.

However, interesting properties can be derived by the structures described herein concerning the role of the pyrazole nitrogen.

Pyrazole protruding inhibitors: no N-anchoring



Pyrazole twisted inhibitors : N-anchoring



Pyrrrole twisted inhibitors: no N-anchoring

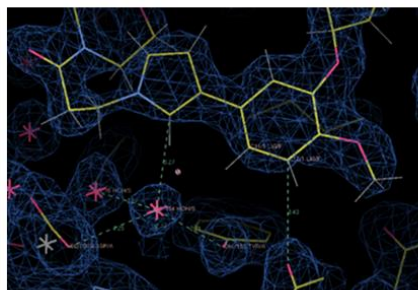


Figure 52: Differences between pyrazole and pyrrole compounds.

Within pyrazole inhibitors, we can appreciate a substantial difference in the role of the “free” nitrogen, between protruding and twisted compounds. While in protruding compounds the nitrogen points outwards towards the entrance of the pocket, in twisted compounds it is oriented towards the bottom of the pocket, forming a water-mediated interaction with residues Tyr325 and Asp482. As already pointed out, this interactions have been also exploited by other PDE4 important inhibitors such as roflumilast ¹⁰⁴.

In our structural investigation, we described that in twisted pyrrole compounds, the lack of the free nitrogen leads to an increase in the distance of 0.3 Å between the position of the former nitrogen (now carbon) and the hub water molecule.

The analysis of this sub-class of inhibitors, together with all the considerations on the conformational variability of the tails, led us to the design of new compounds.

3.6 Structure-based design of novel compounds

3.6.1 Imidazole compounds

The idea of replacing the pyrazole moiety with an imidazole stems from two main considerations:

- The pyrazole compounds that feature a protruding tail are not able to exploit the interaction with the hub water molecule located at the bottom of the catalytic pocket. On the contrary, twisted pyrazole compounds do exploit this interaction.
- Protruding compounds are the most interesting group of inhibitors as they are generally more potent and their tails develop towards a position (the S-pocket) where an opportunity to interact with the UCR2 domain may be provided.

Based on these two considerations, we designed a new class of inhibitors that carry a protruding tail and an imidazole heterocycle in the linker region. In this way, they should be able to exploit the characteristic associated to a protruding tail and the interaction with the lower pocket.

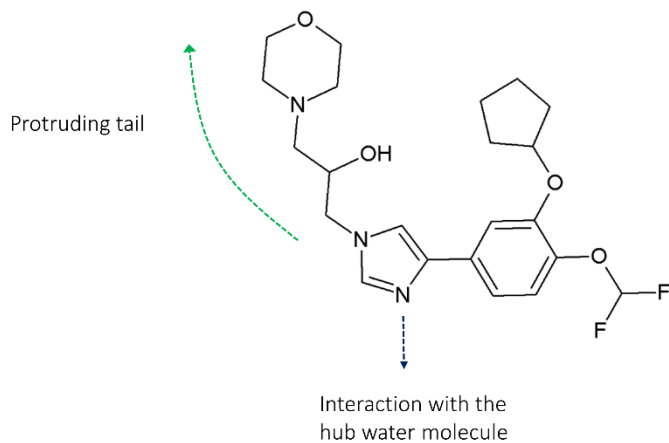


Figure 53: Prototype molecule - GEBR-PP1

This new class of inhibitors is currently being synthesized in the Bruno's lab at the University of Genova.

3.6.2 Condensed cycles

Changing the main scaffold of the compound can be another way to optimize the interaction with the highly polar groove that is located at the bottom of the catalytic pocket.

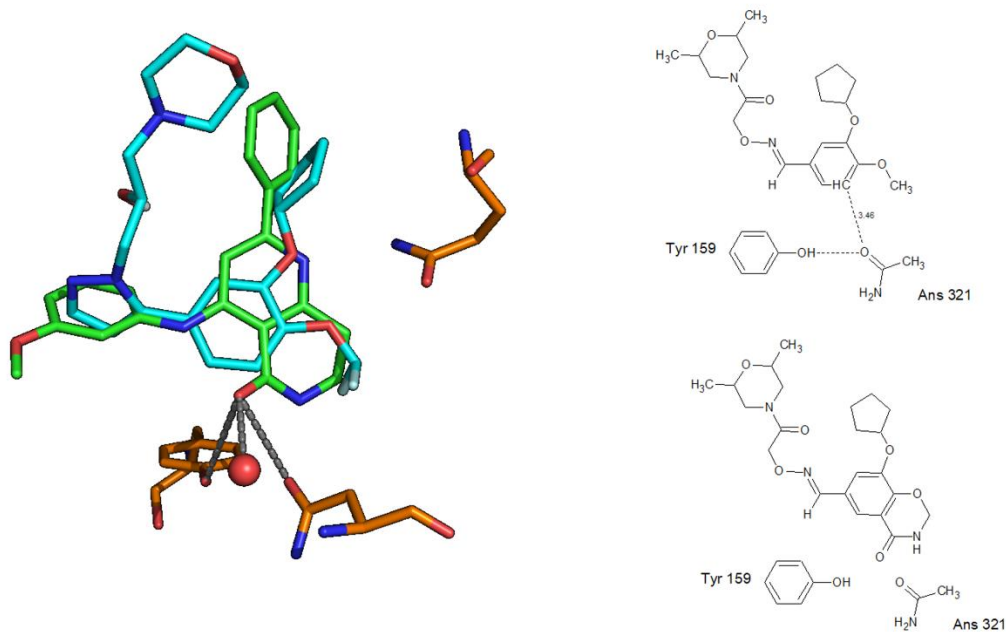


Figure 54: (Left) Superimposition between the structures of PDE4 in complex with GEBR-32a (cyan molecule - PDB ID: 6FDC) and ligand 6Q2 (green molecule - PDB ID: 5K32). (right) schematic view of the lacking interaction in a GEBR-inhibitor binding (top) and in the proposed chimeric molecule GEBR-PP2 (bottom)

The differences between GEBR-32a and compound 6Q2¹⁰⁵ bound to the PDE4D catalytic domain are shown in (Figure 54). The compound 6Q2 is able to anchor itself to the bottom of the pocket by forming multiple hydrogen bonds via its amide moiety. Conversely, the catechol of GEBR-32 does not feature any functional group that can act in the same way. Based on our structural data,

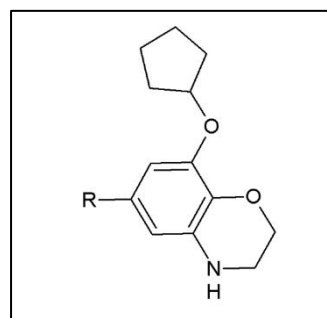


Figure 55: GEBR-PP3 central scaffold

we therefore propose to synthesize a novel compound where a condensed cycle attached to the catecholic portion is introduced. The new scaffold maintains the two oxygen atoms that are required for the Q-switch interaction while the new portion is developed towards the inner pocket, where an amide moiety could exploit a new set of polar interactions, similarly to compound 6Q2.

Unfortunately, the proposed compounds can be hardly synthesized. Instead, a possible analogue could be GEBR-PP3 (see Figure 55), which features an amine instead of an amide. GEBR-PP3 is currently being synthesized in the Bruno's lab at the University of Genova.

3.6.3 Tail tuning

Our X-ray crystallography studies have highlighted the ability of the protruding compounds to explore the S-pocket, pointing towards the external portion of the catalytic cavity, in the proximity of the area that in long isoforms is occupied by the UCR2 helix. As already mentioned, this structural property is quite innovative for PDE4 inhibitors, and only a few other recently developed compounds are able to exploit this space. However, GEBR compounds did not show a clear differential potency in the inhibition of the PDE4 catalytic domain and the PDE4D3 long isoform. Our results suggest that a fine-tuning of the chemical properties of the tails could be a strategy to maintain the orientation of the compounds while optimizing the possibility to interact with the regulatory capping portion. In Figure 56, the structure of an allosteric inhibitor³¹ (grey) that is able to stabilize UCR2 capping by Phe196 clamping (PDB ID: 3IAD) is shown, superimposed with the structure of GEBR-54 (purple).

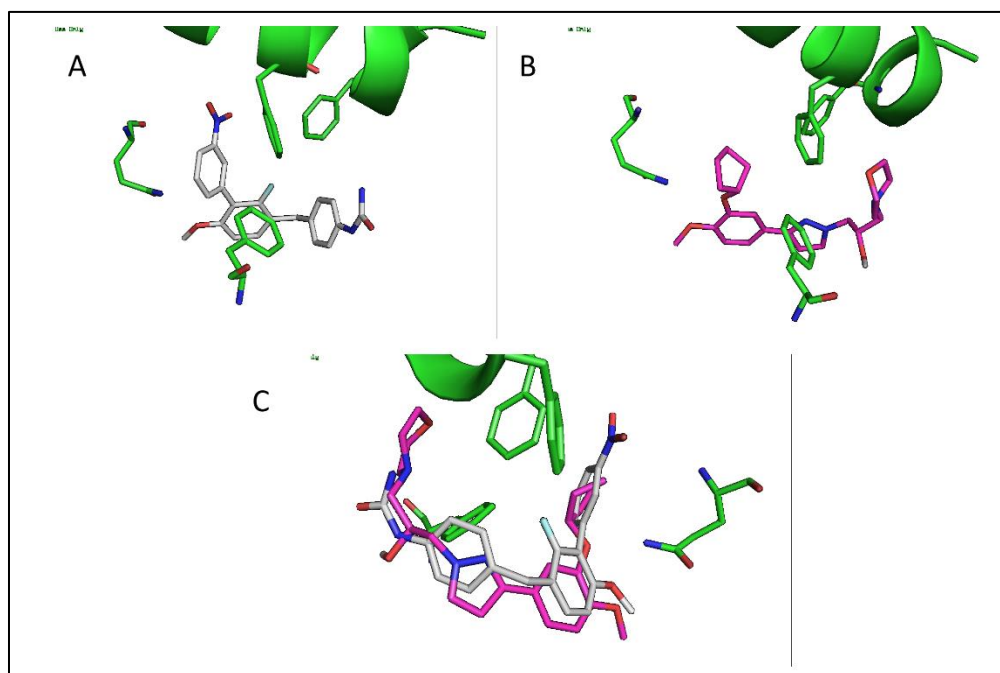


Figure 56: A) details of the binding pose of a PDE4 allosteric modulator. B) details of the binding pose of GEBR-54. C) superposition between the two structures

While the allosteric modulator is perfectly shaped for nesting two phenylalanine residues of the UCR2, GEBR-54 could form a groove under Phe196. The general rationale behind our prototype molecules is to exploit slightly shorter tails (more similar to NAMs) with an increased hydrophobic nature that could fit better with the hydrophobic lid of UCR2 (Phe196 and Phe201). Moreover, the

substitution of the pyrazole with an imidazole could serve also to improve the hydrophobicity of the groove where Phe196 will insert. The molecules proposed are currently being evaluated for their synthetic feasibility.

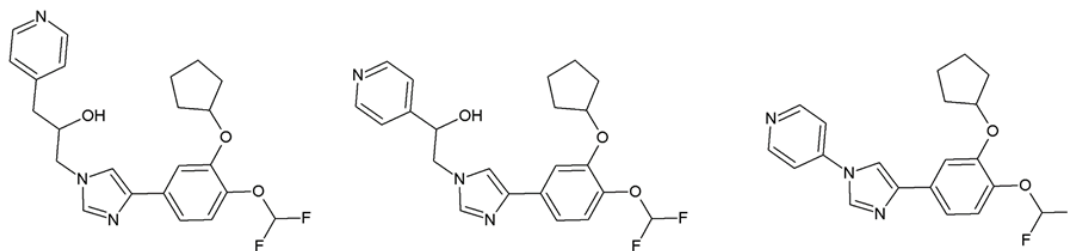


Figure 57: Set of compounds derived from GEBR-32a that feature 1) a progressively shorter tail, 2) an aromatic ring instead of a morpholine and 3) a central imidazole-based linker

4 Conclusions

To contribute to the elucidation of the molecular bases of selective inhibition of different PDE isoforms, I described the structural and functional characterization of a number of members of the GEBR family of PDE4D inhibitors. Indeed, when tested in animal models, some of these compounds have been previously shown to cause limited side effects relative to other PDE4 inhibitors, suggesting some degree of specificity or lack of cross-reactivity among different PDE4 isoforms. Within all the crystal structures of the PDE4 catalytic domain shown herein, each in complex with a different member of the GEBR library of compounds, three major ligand conformational classes were identified and we classified them as protruding, twisted and extended. Whereas all the ligands display a common catechol portion that binds into the PDE catalytic pocket similarly to Rolipram, they also feature different chemical scaffolds of variable lengths, thus possibly allowing some degree of interaction with the regulatory domains of the enzyme. Owing to the known intrinsic difficulty in obtaining the crystal structure of the full length enzyme, we assessed the ability of these compounds to interact with portions of the enzyme outside the catalytic pocket by examining their binding conformation inside the active site of the catalytic domain and by comparing their potency measured against the full length protein and the catalytic domain only. To further validate our conclusions based on these sets of observations, molecular dynamic simulations were used to generate a model of the full length protein in complex with two of our ligands, each representative of the two general chemical scaffolds based on which they had been designed. Even though only weak and unspecific hydrophobic interactions between the tail portion of the ligand and the capping portion of the enzyme (Phe196) can be postulated based on our simulations, which is also in agreement with the lack of a large differential inhibitory behavior, the conformations of both ligands that were derived from our crystal structure analysis were maintained over the course of the molecular dynamics simulations, suggesting a certain degree of stability of the complex.

Based on the data shown herein and on the pharmacological profiles of selected GEBR library compounds previously described in the literature, the subset of protruding ligands appears to be the most promising candidate for further library development. Within this conformational ensemble, the shorter compounds (GEBR-7b and GEBR-20b) do not inhibit differentially the PDE4 catalytic

domain and PDE4D3, while with the lengthening and the enhanced flexibility of the tail a differential behavior starts to appear (GEBR-4a and GEBR-11b). Moreover, within the protruding compounds, the one with the greatest differential behavior is GEBR-54. Interestingly, the GEBR-54 structure shows greater flexibility than those featuring the standard protruding conformation in the S pocket.

In April 2018 the project reached its first milestone, with the publication in the journal *ACS Biochemistry* of the structural and the functional characterization of the most interesting GEBR library inhibitors, which are presented in this PhD thesis.

Currently, further research activities are ongoing and the project is currently divided in two lines:

- GEBR-PPs design, feasibility assessment and synthesis. Our characterization yielded an interesting amount of information that are being used to design new inhibitors (see chapter 3.6). Imidazole-based compounds (GEBR-PP1) and condensed cycles (GEBR-PP3) are now being synthesized.
- Characterization of other GEBR library inhibitors. The data shown in this thesis relates to a small selection of the GEBR-library compounds, with particular attention to those inhibitors that have also been studied in animal models. However, a huge portion of the library remains uncharacterized.

Structural insights into the interplay between cis- and trans-interactions in Cadherin-mediated cell adhesion

5 Introduction

At the adherens junctions, cell-cell interactions are mostly mediated by a large family of calcium-dependent transmembrane proteins called cadherins. Two layers of cadherins that protrude from two adjacent cells interact with each other to form an adhesive interface resembling a molecular Velcro. Collectively, the mechanistic behavior of the system determines the adhesive phenotype of a cell. However, the cellular role of cadherins is not limited to forming an anchoring interface for tissue morphogenesis and stability. Indeed, cadherins act also as mechano-transducing molecules,

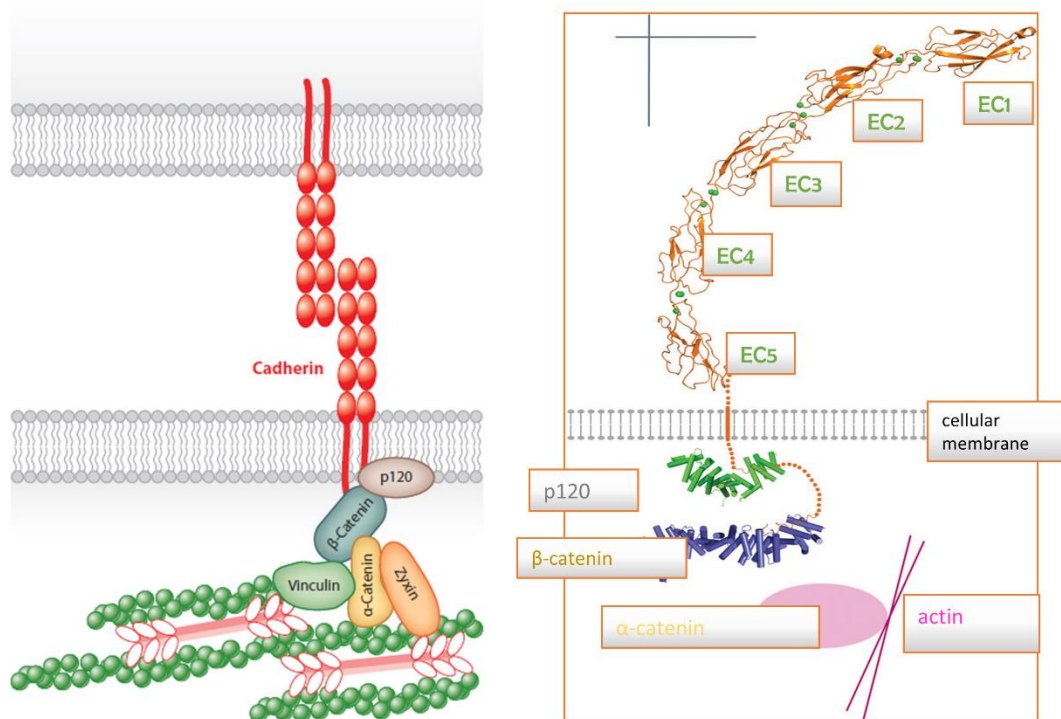


Figure 58: Architecture of cadherin-mediated junctions. Left: schematic view From Leckband and Rooij. *Annu. Rev. Cell. Dev. Biol.* 2014. 30:291-315. Right: schematic view merged with PDBs structures, from Brash at al, *Trends in cell Biology* 2012

sensing changes in tension at the adherents junction¹⁰⁶ and taking part in the signaling process involved in cytoskeletal organization, cell cycle progression and differentiation^{106,107}.

Cadherins are a large and phylogenetically diverse superfamily of cell adhesion molecules. Among all the members of the cadherin superfamily, the so-called *classical cadherins* subfamily is historically the most studied and characterized. The overall architecture of classical cadherins features three main regions:

- Extracellular portion: it is organized in five immunoglobulin-like domains (EC1-EC5, where EC1 is the outmost domain and EC5 is the membrane proximal domain) arranged in tandem and it provides the protein with its adhesive properties. Adhesion takes place mostly at the level of the last two external domains.
- Transmembrane portion: it is a single pass helix that crosses the plasma membrane
- Cytoplasmic portion: it is a highly conserved intracellular domain that features a catenin (α , β , γ) binding site. The interaction with catenin molecules allows communication between the cadherin extracellular portion and the actin cytoskeleton. The cadherin cytoplasmic domain is completely unstructured in the absence of binding partners.

5.1 Mechanism of cadherin dimerization

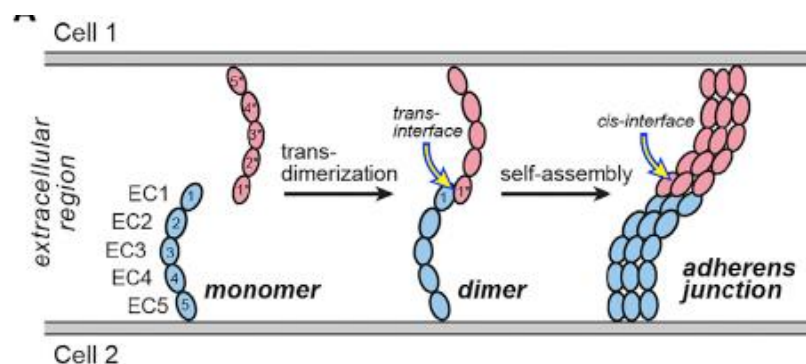


Figure 59 : Cadherin-mediated adhesion – generation of a junction.
From S. Kudo et al. *Structure* 2016. 24: 1523-1536

Within the adherence junction, we recognize the single interaction between two cadherins protruding from two neighboring cells as the fundamental building block of a fully adhesive

interface. Such binding event occurs via a mutual exchange of the six aminoacid-long N-terminal portion of the EC1 domain, the so-called adhesion arm. Two cadherins that have completed their



Figure 60: P-cadherin monomer. PDB ID: 4OY9 (left) and P-cadherin final strand-swap. PDB ID: 4ZML. Adhesion arms are enlighten in yellow

dimerization trajectory form a homo-dimeric complex whose conformation is usually referred to as the “strand-swap” dimer^{106–108}.

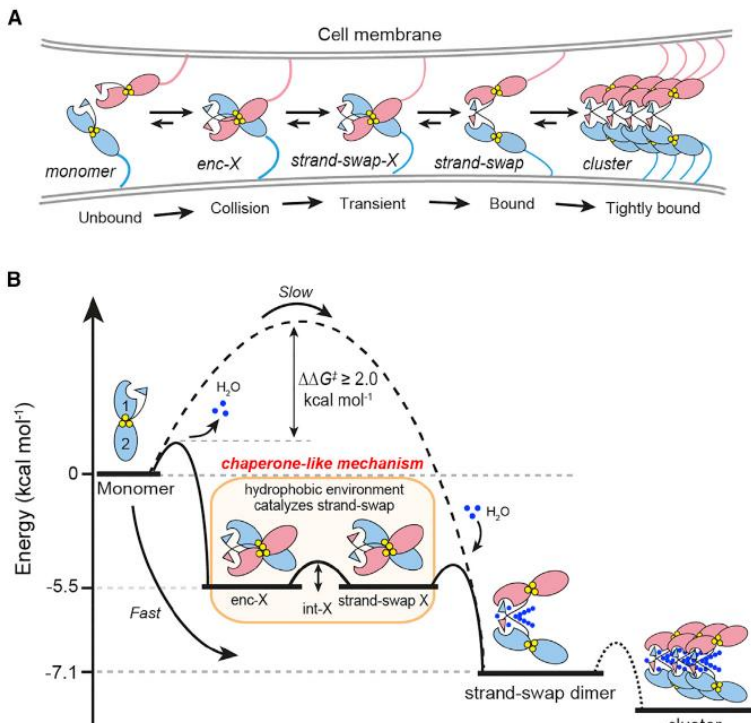


Figure 61: Cadherin adhesion mechanism scheme

Over the years, mechanistic studies on cadherin dimerization have shown that the dimerization pathway that leads from monomeric cadherin to the strand swap dimer goes through a series of crucial intermediates, some of which have been fully characterized at the structural level¹⁰⁹:

1. Monomeric cadherin: for a long time, this form has been considered only related to the non-activated version of the protein (featuring the N-terminal pro-domain). Parisini *et al.* demonstrated

that monomeric P-cadherin actually exists after its priming (pro-domain removal). Moreover, its adhesive arm fluctuates constantly in and out the binding pocket, until the system is able to proceed along its dimerization pathway¹¹⁰.

2. In-between states: The recognition between two cadherins starts with weak interactions at the level of the calcium domains. This weakly adhesive dimeric arrangement, called the “X-dimer”, has been shown to encompass a family of conformations ranging from the so-called first encounter X-dimer complex (enc-X) to the initial exchange of the adhesive arms (SS-X). From an energetic point of view, this checkpoint is a meta-stable conformation that lowers the energy barrier associated with the transition from the monomer to the strand-swap dimer. This kinetic advantage is driven by the formation of a non-polar environment at the level of the N-terminal swapping regions.
3. Strand-swap: The final step of cadherin dimerization occurs upon arms exchange and gradual separation of the EC2 domains.

The conformational pathway along which the cadherin homo-dimerization process occurs is characterized by an extremely smooth potential energy surface that accelerates both association and dissociation. The direct consequence of this reversibility is the dynamic behavior of the adhesion surface and its continuous remodeling.

5.2 The interplay between the *trans* and the *cis* interface.

As already pointed out, the association between two cadherins protruding from two adjacent cells can be considered as the minimal building-block for the formation of an adhesive junction.

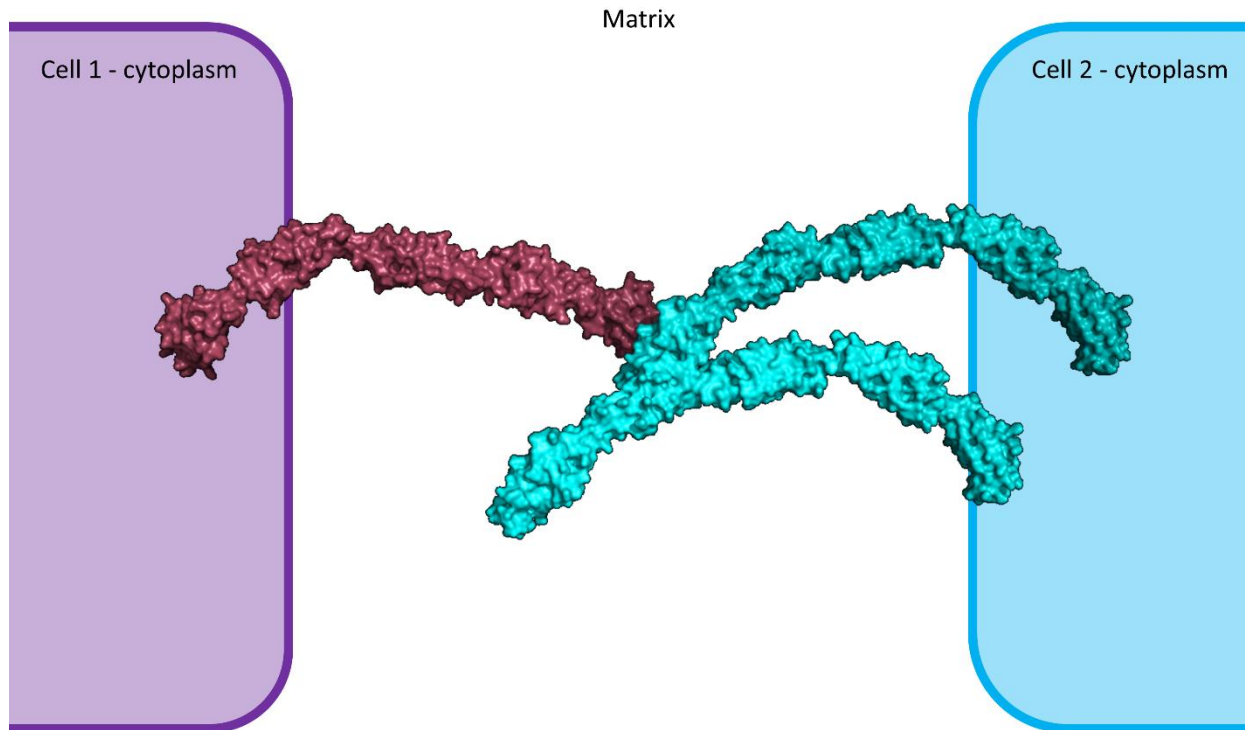


Figure 62: Schematic representation of the *cis*-(two cyan cadherins protruding from the same cell) and the *trans*-(cyan and magenta cadherins protruding from opposite cells) interaction. Protein structure from PDB ID: 1LW3

However, adherens junctions are large cellular surface areas hosting a multitude of identical cadherins acting cooperatively to provide sufficient strength to hold two adhering cells together. Electron microscopy studies that have elucidated the “Velcro-like” nature of the full junction supported the notion, previously derived by numerous crystallographic studies, that also interactions that are formed between cadherins protruding from the same cell may play a fundamental role for the formation of the junction, as well as for its stability.

The packing arrangement in the crystal structure of the full ectodomains of both C- and E-cadherin revealed a conserved interaction between the EC1 of one protein and the EC2 of the adjacent protein. In the crystal lattice, such proteins are oriented as if they are ideally protruding from the same cell. This is the so-called “*cis*” interaction. It involves an epitope of the EC1 that is located on the opposite side relative to the adhesion arm (i.e. where the EC1 is involved in forming the

“*trans*” interaction) and the apex of the EC2 of the adjacent protein (Figure 62). The organization of the lattice thus produced in the crystal provides an “ideal” yet partially realistic representation of an adherens junction between two cells. The *cis* interaction is mostly hydrophobic in nature. Therefore, owing to the weak stabilizing energy of the *cis* interaction and to the difficulty to study it by binding experiments, its biological role has over the years been investigated only by cell-based and liposome-based experimental approaches on WT/mutant cadherins, yielding some interesting insights into the cadherin activation mechanism. Indeed, mutations that impaired the *cis* interaction were shown to:

- produce reduced levels of adhesion in liposomes
- cause the junction to become unstable in cells expressing also WT cadherins
- impair the ability of cadherins to clusterize, i.e. to undergo a self-assembly process that is of fundamental importance for junction formation.

Albeit inconclusive, these studies provided clear evidence not only of the importance of the *cis* interaction, but also of its possible interplay with the *trans* interaction. The modern model of junction formation describes a process in which the lateral packing of cadherins occurs upon *trans* dimerization. This is in accordance with some observations that cadherins are unable to cluster in the absence of a partner cell eliciting adhesion^{111,112}.

5.3 Cadherins as “undruggable” targets

The adhesive properties of a cell are responsible for (or connected with) polarity, differentiation, growth and many other cellular mechanisms. In tumors, the adhesive phenotype is highly disturbed by genetic and epigenetic factors, giving rise to aberrant properties such as for instance altered cell migration. The involvement of cadherins in cancer has been extensively reviewed by Berx and van Roy¹¹³. In Table 3 some key facts about E and P-cadherin are reported.

Table 3: Data extrapolated from Berx and van Roy, CSH Perspective 2009

<i>Protein</i>	<i>Molecular event</i>	<i>Tumor type</i>	<i>Correlation</i>
E-cadherin	Promoter methylation Germline mutations	Various	Malignant progression

		Gastric (DGC), Breast (ILC)	Hereditary gastric cancer syndrome
	Somatic mutations	Gastric (DGC), Breast (ILC), Pancreas	Highly invasive grown pattern
	Up-regulation	Epithelial Ovarian cancer	Tumorigenesis
	Up-regulation	Breast	Tumor emboli formation
P-cadherin	Up-regulation (cadherin switch)	Breast, Gastric, Pancreas (PDAC)	Enhanced migration and invasion
	Down-regulation	Melanoma	Increased invasion and metastasis

Considering their roles in tumorigenesis, cadherins are now considered as possible targets for cancer therapy. In principle, future anti-cancer drugs should be tailored to be more effective or selective, responding to specific abnormalities of the adhesive phenotype¹¹³.

The development of small molecules that are able to inhibit or to modulate cadherin-mediated cell adhesion has been proposed as a possible strategy to specifically sensitize tumors to chemotherapeutic agents. The first successful example is ADH-1, a cyclic pentapeptide that contains the recognition

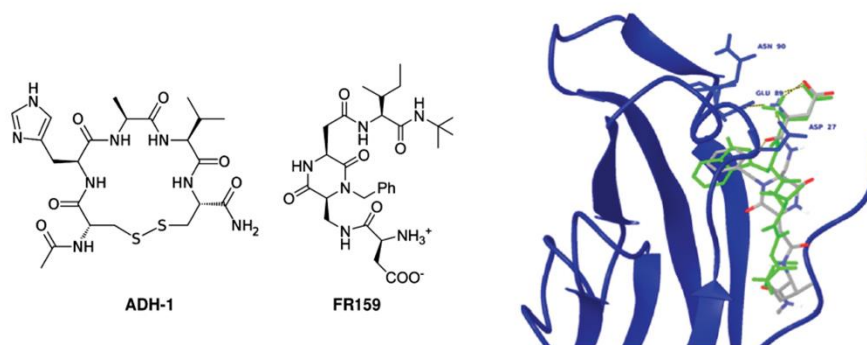


Figure 63: Left - ADH1 and FR159 compounds. Right - FR159 computationally-derived binding pose. From Doro et al. *Organic & Biomolecular Chemistry Communication* 2015.

sequence His-Ala-Val and that has been shown to inhibit N-cadherin adhesion in solid tumors. This drug candidate has been studied in phase 1 clinical trials, yielding only modest results¹¹⁴.

A library of peptidomimetic compounds has been developed by Doro and colleagues with the aim of mimicking the E/N-cadherin adhesion arm (DWVI sequence). These peptidomimetics feature a

phenyl ring in substitution of the indole moiety that constitutes the side chain of the Trp2 residue that is crucially involved in the strand-exchange cadherin homo-dimerization mechanism. Within the set of compounds that have been developed in this library, FR159 resulted to inhibit N-cadherin-mediated adhesion in EOC cells, with an increased potency relative to ADH-1 (1mM)¹¹⁵. Interestingly, high resolution structural investigation by means of X-ray crystallography provided clear evidence that the actual binding region of FR159 differs from the pose that had been postulated based on computational methods¹¹⁶. Indeed, the inhibitor binds across two interacting cadherins at the level of the hydrophobic pocket that is generated upon X-dimer formation and lined with the residues Ile4, Pro5, Ile7 and Val22 of both cadherins. Unexpectedly, the crystal structure of the complex showed the existence of a newly identified pocket, which can now be targeted with a structure-based drug-design approach in future studies. The crystal structure shows that FR159 binds the cadherins in the X-dimer conformation in a position that precludes the system

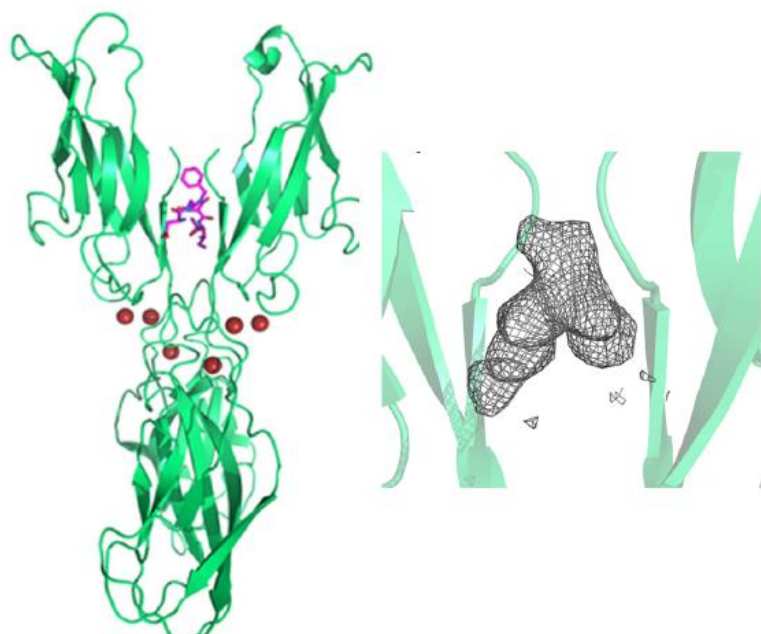


Figure 64. Left – crystal structure of E-cadherin-(V3) in complex with FR159. Right – Electron density map of FR159 in the same structure.. (PDB ID: 4ZTE).

to evolve towards the final strand-swapped state¹¹⁶.

5.4 Aim of the project

Although the notion of a strictly conserved mutual interplay between the *cis* and the *trans* interaction in classical cadherins has long been accepted within the scientific community and

partially confirmed using cell adhesion assays, no actual experimental validation of this postulate has ever been reported. Mechanistic investigations of cadherin-mediated cell adhesion and its modulation are a core research line in the Parisini's lab. In this framework, we set out to study by X-ray crystallography the conformational behavior that cadherins adopt when key mutations are introduced to perturb the interaction surfaces that are fundamental for its progression towards the strand-swap dimer conformation. P-cadherin is the sole member of the family that is able to crystallize in its monomeric form when completely primed for adhesion (proteolytic removal of the N-terminal pro-domain). Indeed, the existing crystal structures of the closed forms of E- and N-cadherin have been obtained by engineering mutants where one or more extra residues had been inserted at the N-terminus of the protein, thus providing a bulky N-terminus that, owing to its steric hindrance, would not be able to progress towards the strand-swap dimer conformation. This unique feature makes P-cadherin the perfect system to study all the metastable conformations that can be explored, and therefore perturbed, during dimerization. In fact, in crystallization experiments selected and carefully designed P-cadherin mutants can in principle adopt spontaneously any of the three major possible cadherin conformations (monomer, X-dimer, strand-swap dimer). It is therefore in such system that it is possible to investigate the mutual relationship between the *trans* and the *cis* interaction. Ideally, we aim to introduce minimal-impacting mutations that are able to modify the potential energy surface and that can let us explore each possible metastable potential energy minimum of the system (monomer, X-dimer, strand-and monomer again) and address the interplay between the *cis* and the *trans* interaction.

6 Materials and methods

6.1 Constructs

All protein mutants were generated by Agilent Quick-change II site directed mutagenesis kit, starting from a clone of P-cadherin-EC1-EC2 A5P, cloned in pET3a. Each construct maintained the 6xHis tag at the N-terminal followed by a spacer and the enterokinase cleavage region (DDDDK).

Starting template P-cadherin-EC1-EC2 A5P sequence:

```
gattgggtggttcccctatcagtgtgccg
      D W V V P P I S V P
gagaacggcaaaggtccctttccgcaacgccttaaccaactgaagtccaacaagaccgt
      E N G K G P F P Q R L N Q L K S N K D R
gatacgaaaatcttctacagtattacaggacctgggtgctgacagccctccggaaggcgtg
      D T K I F Y S I T G P G A D S P P E G V
tttgcggttgagaaagaacgggttggttgctgcttaacaaccctcgatcgcaagaa
      F A V E K E T G W L L L N K P L D R E E
attgcgaaatacagagctgtttggtcatgcggtttcggaaaatggcgcctcagttgaagat
      I A K Y E L F G H A V S E N G A S V E D
ccgatgaacatcagcatcattgtcacagaccagaacgaccataaacgaaatttaccag
      P M N I S I I V T D Q N D H K P K F T Q
gatacatttcgcggtagtgtgctggaaggggtcctcccaggaacctcagtgatgcaagtt
      D T F R G S V L E G V L P G T S V M Q V
actgcgaccgatgaagatgatgcatctacacgtataacggtgtgtagcatactcatt
      T A T D E D D A I Y T Y N G V V A Y S I
cactcgcaagaacccaaagatccgcatgacctgatgttcacgattcatcgctctacaggt
```

H S Q E P K D P H D L M F T I H R S T G
 accatttcagtgatttcgagcggtttagaccgtgagaaagtgccggaatatacgtgacg
 T I S V I S S G L D R E K V P E Y T L T
 atccaggctactgacatggacggcgatgggagtactaccaccgcagttgccgtttagag
 I Q A T D M D G D G S T T T A V A V V E
 atcctggattga
 I L D -

Primers for generating the human P-cadherin-EC1-EC2 A5P P123G mutant:

- Forward primer: ctggaaggggtcctc**gga**ggaacctcagtgatg
- Reverse primer: catcactgaggttc**cc**gaggacccttcag

Primers for generating the human P-cadherin-EC1-EC2 A4P P123G K14E mutant:

- Forward primer: gtgtgccggagaacggc**gaa**ggtcctttccgcaacg;
- Reverse primer: cgttgcggaaggac**ttc**gccgttctccggcacac;

Protocol for PCR:

5µl of 10X reaction buffer

- 125 ng circa of template plasmid DNA (1-5 µL)
- 1µl of forward primer (125 ng)
- 1µl of reverse primer (125 ng)
- 1µl of dNTPs
- 41µl of water
- 1µl of HF DNA polymerase (2.5 U)

The thermal cycle:

- Initial denaturation step: 95°C for 30 seconds
- PCR cycles were repeated 18 times and were composed as follows:
 - Denaturation: 95°C for 30 seconds
 - Annealing: 60°C for 1 minute
 - Extension: 60°C for 6 minutes, since the rate of transcription of transcription of the DNA polymerase is 1kb/min.
- Final extension for unpolymerized zones: 68°C for 7minutes

After the PCR, methylated parental DNA was removed by adding 1µl of DpnI and incubating at 37°C for 1 hour.

The product obtained was used to transform XL1B competent cells with the standard protocol (see materials and methods - Chapter 2). A single colony was selected, grown overnight and the plasmid DNA was extracted by standard techniques (see material and methods – chapter 2). The constructs were finally validated by sequencing (in outsourcing to myGATC).

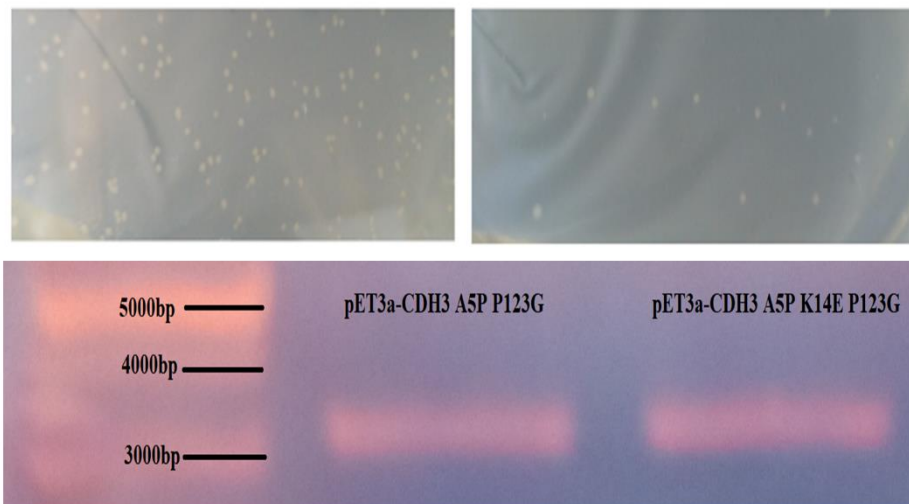


Figure 65: Upper panel - colonies formed upon transformation of XL1B cells with the mutagenic PCR product. Bottom panel - Electrophoresis control after the miniprep

6.2 Protein expression and purification

The plasmids thus created were transformed into *E.coli* BL21(DE3) pLysS cells (Invitrogen) for protein heterologous expression. The cells were cultured at 37° in Luria-Bertani (LB) broth

supplemented with 50 mg/L ampicillin, until OD₆₀₀ reached 0.6. Expression was induced with 1-thio-β-D-galactopyranoside (IPTG, final concentration: 0.8 mM), and the culture was left on an orbital shaker at 25°C overnight. The downstream process involved 3 phases:

1. Preliminary purification: Ni-NTA affinity chromatography and size exclusion chromatography
2. Cadherin activation: proteolytic cleavage of the N-terminal tag (6xHistag-linker-cleavage site)
3. Separation of the activated proteins: refolding (if necessary), counter Ni-NTA affinity chromatography and size exclusion chromatography

6.2.1 Preliminary purification

The general protein buffer is standardized to TBS pH 7.4 + 2 mM CaCl₂. Supplements to lysis buffer and Ni-NTA affinity chromatography buffers, as well as the operative procedures, are equal to the ones used for the purification of the PDE4 catalytic domain described previously (see paragraph 2.1).

Ni-NTA affinity chromatography:

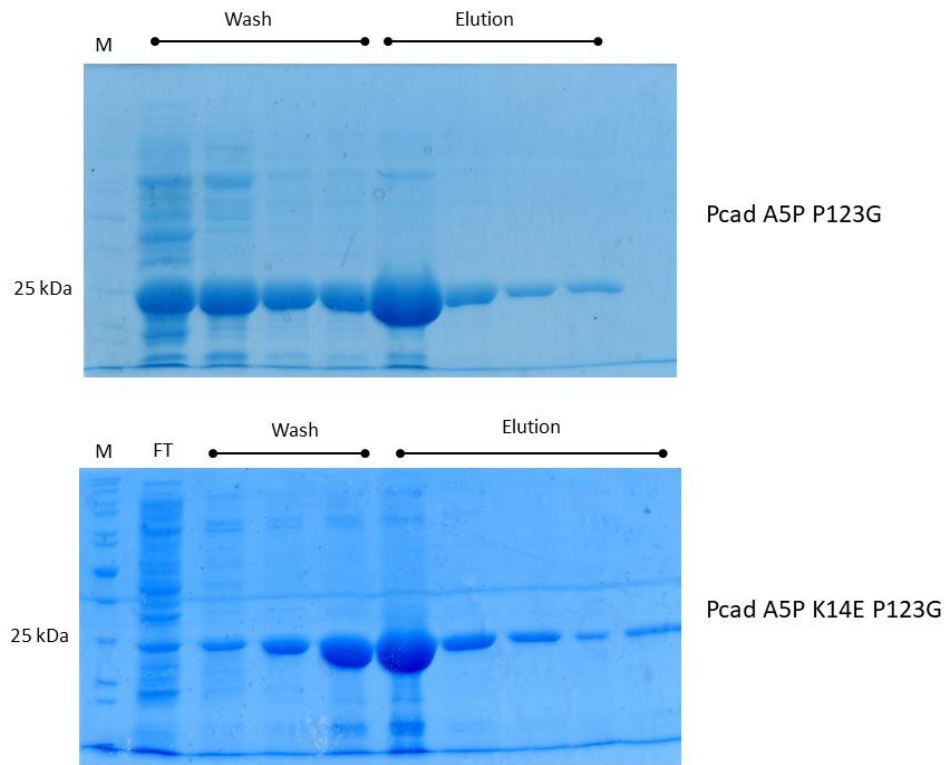


Figure 66: SDS-pages of Ni-NTA affinity chromatography

Size exclusion chromatography:

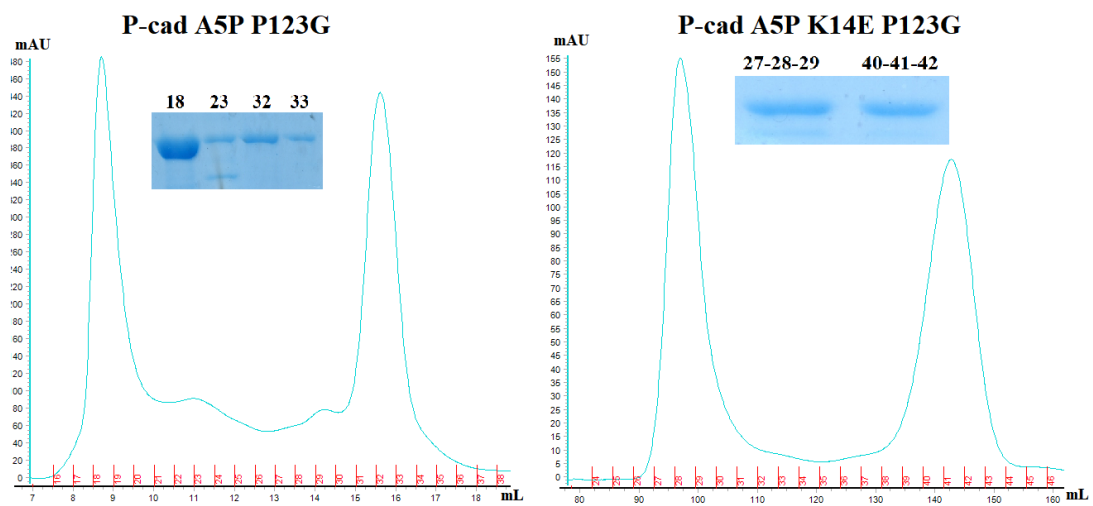


Figure 67: Size exclusion chromatograms and SDS-pages

6.2.2 Cadherin activation:

Enterokinase digestion was used to remove the His-tag extension from the N-terminus of the protein following the protocol:

1. Concentrate protein sample up to 2 mg/ml
2. Adjust CaCl₂ concentration to 20 mM
3. Add enterokinase light chain (NEB – P8070L) to a final concentration of 2 μg/mL
4. Incubate for 72 hours, at room temperature, over an orbital shaker

6.2.3 Separation of the activated cadherins

The enterokinase reaction often resulted in protein aggregation. Whenever needed, a refolding step was introduced in order to recover, at least partially, the aggregated sample.

Refolding procedure:

1. After the enterokinase reaction (approximately 10 ml of total volume), spin-down the aggregated portion (10.000 rpm, 5 minutes)
2. Resuspend the pellet in denaturation buffer (TBS pH 7.5, 20 mM CaCl₂, 6M guanidinium hydrochloride)
3. Dialysis (1) in 1L of TBS pH 7.5, CaCl₂ 5mM, arginine 220 mM
4. Dialysis (2) in 1L of TBS pH 7.5, CaCl₂ 2mM, arginine 100 mM

Counter Ni-NTA chromatography – Size exclusion:

Ni-NTA affinity chromatography was utilized for the removal of the cleaved tag. The protein sample was incubated in the Ni-NTA charged gravity column and then the flow-through was collected. The eluted sample was then further purified by size exclusion chromatography and finally concentrated to 10 mg/ml for crystallization experiments.

6.3 X-ray crystallography

For both the P-cadherin mutants described herein, crystallization experiments were conducted by the hanging drop technique, using VDXm 24-well plates (Hampton research), mixing in each well 1 μ L of protein solution with an equal volume of crystallization buffer, and equilibrating the system against either 0.5 or 1 ml crystallization buffer in the reservoir. Screening and subsequent rounds of optimization were conducted starting from the previously identified crystallization conditions¹¹⁰:

- Polyethylene glycol 8kDa or Polyethylene glycol 12kDa, from 10 to 20%
- 0.1 M HEPES pH 7.5
- 0.2 M CaCl₂

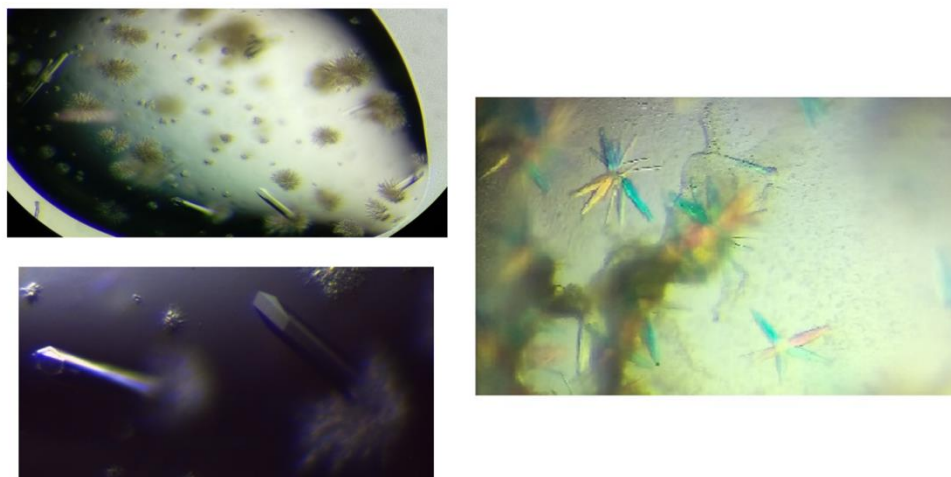


Figure 68: Crystals of P-cad A5P P123G (left) and A5P P123 K14E (right)

- DMSO 0/5/10%

Crystals were cryo-protected using the reservoir solution supplemented with 30% glycerol (P-cadherin-EC1-EC2 A5P P123G) or 30% sucrose (P-cadherin-EC1-EC2 A5P P123G K14E), and flash-frozen in liquid nitrogen. X-ray diffraction experiments were carried out at the Swiss Light Source (SLS) synchrotron, Paul Scherrer Institute (PSI), Villigen, Switzerland. Hardware and experiments details:

- Beamline: X06DA
- Detector: PSI Pilatus 6M
- Wavelength: 1 \AA

- Software used for structure solution:
- Integration of diffraction and data merging: CCP4⁸⁸
- Molecular replacement and refinements: Phenix⁸⁹
- Merging of dataset of isomorph crystals: Blend¹¹⁷. Five different datasets were obtained for both P-cadherin-EC1-EC2 A5P P123G and P-cadherin-EC1-EC2 A5P K14E P123G. Datasets were merged in order to increase completeness and multiplicity.

	P-cadherin-EC1- EC2 A5P P123G	P-cadherin-EC1-EC2 A5P K14E P123G
Space group	P 6 3	P 6 5 2 2
Cell dimensions		
<i>a</i> (Å)	180.988	109.464
<i>b</i> (Å)	180.988	109.464
<i>c</i> (Å)	40.609	94.134
Wavelength (Å)	1	1
Resolution (Å)	2.63	2.5
<i>CC1/2</i>	0.993(0.785)	0.993(0.531)
<i>I/σI</i>	9.5(2.2)	12.4(1.7)
Completeness (%)	100.0(99.7)	99.3(100.0)
Multiplicity	28.5(26.1)	113.2(95.6)
No. of reflections	23217	11920
<i>R_{work}/R_{free}</i>	0.2114/0.2559	0.2124/0.2487
No. of atoms:		
<i>Protein</i>	3322	1661
<i>Water</i>	228	88
<i>Ions</i>	12	4
<i>Other</i>	47	36
Average B-factors (Å ²):		
<i>Protein</i>	81.49	48.22
<i>Water</i>	65.51	49.77
r.m.s.d		
<i>Bond lengths</i> (Å)	0.005	0.002
<i>Bond angles</i> (°)	0.924	0.540
Ramachandran		
<i>Most favoured</i> (%)	86.46	95.73
<i>Additional all.</i> (%)	8.08	2.84
<i>Disallowed</i> (%)	5.46	1.42

Figure 69: Crystallographic table. Unpublished results.

7 Results and discussion

Our goal is to investigate the relative stability of all the metastable conformations that cadherins explore during their dimerization process by introducing punctual and low-impacting mutations that modify the equilibrium between all possible different conformational states. Since we are also aiming at assessing whether the interplay between the *cis* and the *trans* interaction is an absolute requirement for the cadherin dimerization mechanism, our mutations are not designed to modify the N-terminal DAW sequence that allows the system to reach the final strand-swap conformation. Indeed, our set of engineered P-cadherin-EC1-EC2 mutants will all maintain their ability to cycle through all three major cadherin conformations along the full dimerization trajectory, thus possibly allowing us to assess whether a *trans* (strand swapped) interaction may still exist upon impairment of the *cis* interaction.

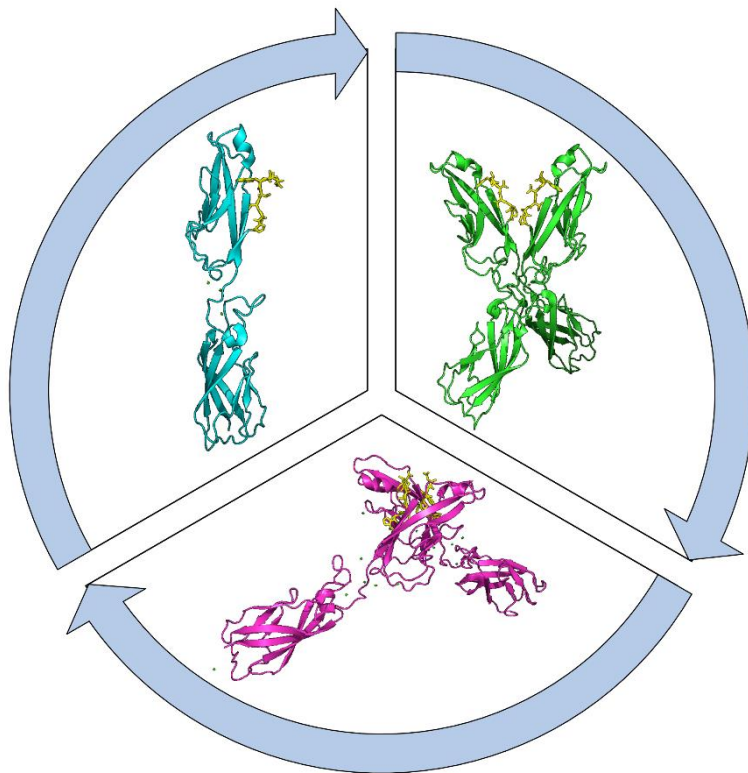


Figure 70: The "life-cycle" of a cadherin molecule primed for adhesion. Monomer (top left), X-dimer (top-right) and strand-swap (bottom) conformation.

We have identified human P-cadherin as the best system for studying the conformational life cycle of cadherins because, unlike all the other classical cadherins studied so far, it crystallizes in its closed monomeric form¹¹⁰ also when primed for adhesion, without the need to lock it in the closed conformation by introducing extra residues at the N-terminus (see Figure 70 and Figure 71). The kinetic differences between E-Cadherin and P-cadherin have been previously investigated by computational methods with the aim to understand the molecular reasons why only P-cadherin is able to retain its monomeric state (in crystals) (unpublished results). The two adhesion arms (DWVVAPIS for human P-cadherin and DWVIPPIS for human E-cadherin) differ for the presence of a diproline motif in position 5 and 6. The presence of two prolines have long been shown to confer an elastic tension outwards, prompting the adhesion arm to open and forcing the protein to crystallize in its strand swap conformation¹¹¹. Indeed, our A5P mutation in human P-cadherin facilitated the opening of the adhesion arm, driving the crystallization towards the strand swap (see Figure 70 and Figure 71).

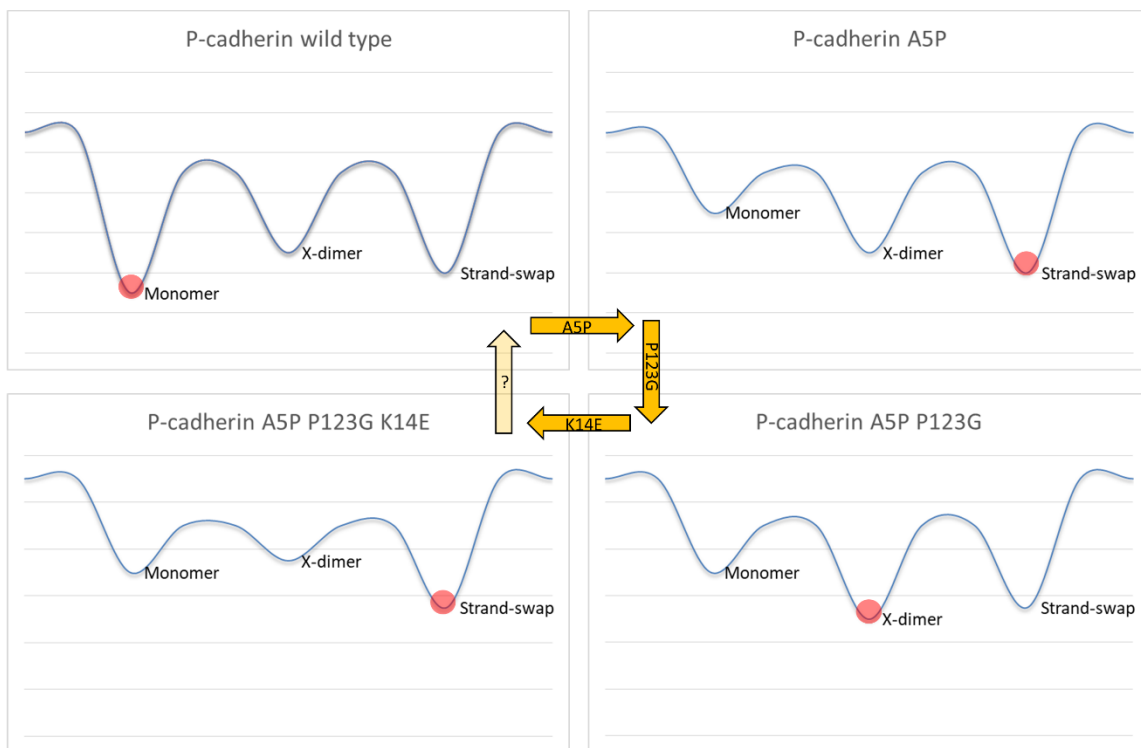


Figure 71: Schematic representation of the potential energy surface of the different P-cadherins constructs (arbitrary scheme).

As previously mentioned, the cis- interaction involves the opposite side of the adhesion arm of an EC1 domain and the apical part of an EC2 domain of a cadherin protruding from the same

membrane. In atomistic details, the two highly conserved regions involved in the *cis* interaction are the HAV (on EC1) and the PGT (on EC2) sequences. At the structural level, the HAV sequence forms a hydrophobic lock in region that accommodates the proline of the interacting PGT sequence. Other hydrophobic residues complete the interaction surface.

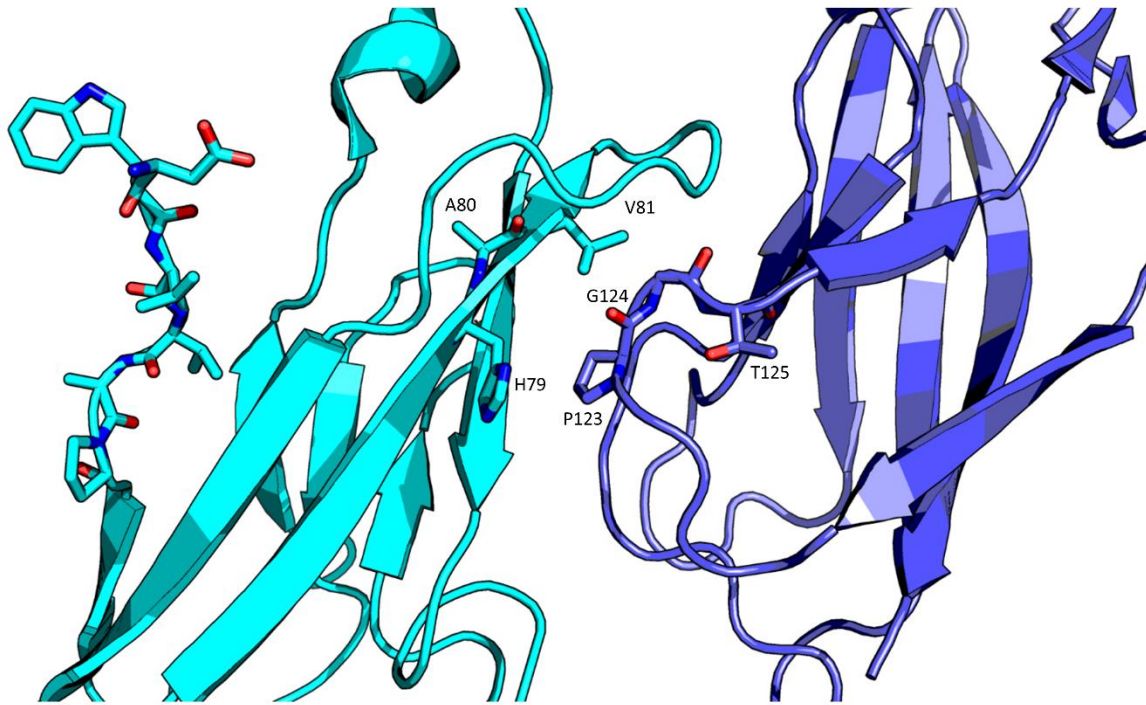


Figure 72: Details of the *cis* interaction between an EC1 domain (left) and the EC2 domain of the adjacent cadherin (right)

The mutation P123G has been designed with the aim of disrupting the *cis* interaction without altering the hydrophobic environment. Indeed, the absence of the bulky side-chain should abolish the key-lock mechanism while conserving the general chemical-physical properties of the surface.

The P-cadherin-EC1-EC2 A5P P123G mutant crystalized in the $P6_3$ space group, adopting an X-dimer conformation. The crystal packing shows two different interactions in the GGT site for the two monomers of the X-dimer:

- EC2 monomer 1: it is sandwiched by two other EC2
- EC2 monomer 2: it is still in close proximity of the EC1 of an opposite cadherin. However, the interaction is disrupted and the relative positions of the two interaction sites (GGT and HAV) are distorted.

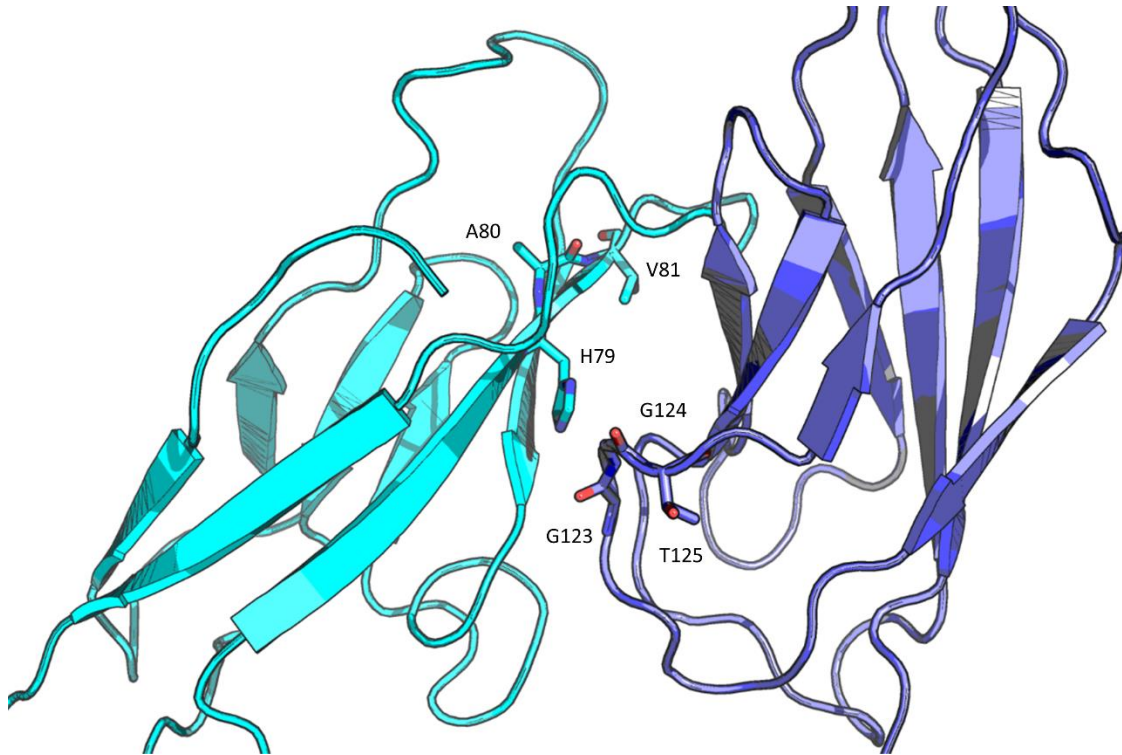


Figure 73: Atomic details of the disrupted *cis*-interaction surface in *P*-cad A5P P123G crystal structure

By impairing the *cis* interaction, the P123G mutation changed the potential energy surface of the dimer, modifying the relative energies of the conformational states and shifting the crystallographic minimum away from the strand swap dimer conformation (see Figure 71). This result supports the

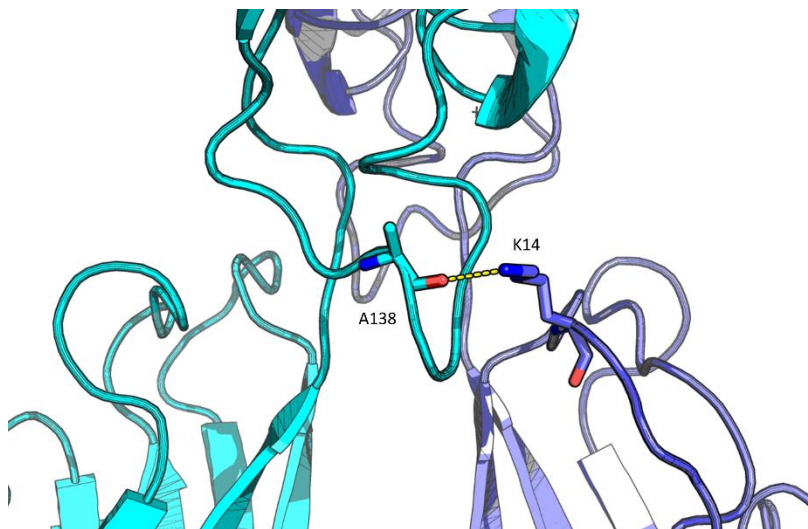


Figure 74: Details of K14-A138 lock in *P*-cadherin X-dimer

notion that the *cis* interaction is crucial for the stabilization of the final strand swap dimer (*trans*) conformation in cadherin-mediated cell-cell adhesion, although it cannot be considered yet a definitive proof that a mutual *cis-trans* dependence is present within the cadherin system.

The K14E mutation is known to hamper X-dimer formation in the crystal. In fact, this mutation prevents the formation of a hydrogen bond with Ala138, which locks

the X-dimer conformation by providing a physical link between the two sides of the adhesion interface. This mutation has been introduced in P-cadherin-EC1-EC2 A5P P123G to assess whether the disruption of the X-dimer could drive the cadherin to adopt again a monomeric conformation in the crystal. Indeed, the disruption of both the *cis* and the X-dimer interaction surfaces could further modify the potential energy surfaces driving P-cadherin to its original monomeric (close) state, thus possibly providing the ultimate proof of the strict interplay existing between the *cis* and the *trans* interaction.

Surprisingly, P-cadherin-EC1-EC2 A5P P123G K14E crystallized as a strand swap dimer. The disruption of the X-dimer interaction generates a new potential energy surface in which the crystallographic minimum is still represented by the endpoint of the dimerization trajectory.

The whole picture provides a qualitative description of the different weights associated to each conformational state, where the *trans* and the X-dimer interactions drive, energetically, the dimerization process, while the *cis* interaction allows a supporting interface to form. The fact that the impairment of both the X-dimer and the *cis* interaction is not sufficient to revert the conformation of the protein to the closed monomer confirms that the A5P mutation confers a completely different dynamic behavior to the adhesion arm, which becomes extremely prone to openings.

Future work will aim at introducing more extreme mutations, in order to further hamper the formation of the *cis* interaction. Indeed, a more invasive modification could extensively or completely abolish the *cis* interaction (for instance, by repulsion). The newly designed mutants will for instance feature those mutations that are known to play a clinical role in different cadherin-associated diseases. The clinical options that we are considering, by looking into the Clinvar database and searching for both the E- and the P-cadherin genes are:

- P123S: E-cadherin germline mutation associated with hereditary diffuse gastric cancer syndrome
- G124R: P-cadherin germline mutation related to ectodermal dysplasia, eterodactyly and macular dystrophy syndrome. E-cadherin somatic mutation associated to invasive breast cancer.

These mechanistic studies, which are still ongoing in the Parisini's lab, will help shed light on those aspects of the complex cadherin homo-dimerization mechanism that are not yet completely understood, thus ultimately facilitating the design of selective cadherin inhibitors for therapeutic applications.

References

1. Krauss, G. *Biochemistry of Signal Transduction and Regulation. The Quarterly Review of Biology* **84**, (Wiley, 2009).
2. Bradshaw, R. A. & Dennis, E. A. *Handbook of Cell Signaling. Handbook of Cell Signaling* **1–3**, (2003).
3. Latorraca, N. R., Venkatakrisnan, A. J. & Dror, R. O. GPCR dynamics: Structures in motion. *Chem. Rev.* **117**, 139–155 (2017).
4. Beavo, J. A. & Brunton, L. L. Cyclic nucleotide research - Still expanding after half a century. *Nat. Rev. Mol. Cell Biol.* **3**, 710–718 (2002).
5. Hurley, J. H. Structure, Mechanism, and Regulation of Mammalian Adenylyl Cyclase. *J. Biol. Chem.* **274**, 7599–7602 (1999).
6. Kleinboelting, S., van den Heuvel, J. & Steegborn, C. Structural analysis of human soluble adenylyl cyclase and crystal structures of its nucleotide complexes-implications for cyclase catalysis and evolution. *FEBS J.* **281**, 4151–4164 (2014).
7. Smith, F. D. *et al.* Local protein kinase A action proceeds through intact holoenzymes. *Science (80-.)*. **356**, 1288–1293 (2017).
8. Sassone-Corsi, P. The Cyclic AMP pathway. *Cold Spring Harb. Perspect. Biol.* **4**, 3–5 (2012).
9. Wong, W. & Scott, J. D. AKAP signalling complexes: focal points in space and time. *Nat. Rev. Mol. Cell Biol.* **5**, pages 959–970 (2004).

10. Taskèn, K. & ... Structure, function, and regulation of human cAMP-dependent protein kinases. *Adv Second Messenger Phosphoprot. Res* **31**, 191–204 (1997).
11. Cheng, X., Ji, Z., Tsalkova, T. & Mei, F. Epac and PKA: A tale of two intracellular cAMP receptors. *Acta Biochim. Biophys. Sin. (Shanghai)*. **40**, 651–662 (2008).
12. Kaupp, U. B. *et al.* Cyclic nucleotide-gated ion channels. *Physiol. Rev.* **82**, 769–824 (2002).
13. Huai, Q., Colicelli, J. & Ke, H. The Crystal Structure of AMP-Bound PDE4 Suggests a Mechanism for Phosphodiesterase Catalysis. *Biochemistry* **42**, 13220–13226 (2003).
14. Wong, K. Y. & Gao, J. Insight into the phosphodiesterase mechanism from combined QM/MM free energy simulations. *FEBS J.* **278**, 2579–2595 (2011).
15. Bolger, G. B., Rodgers, L. & Riggs, M. Differential CNS expression of alternative mRNA isoforms of the mammalian genes encoding cAMP-specific phosphodiesterases. *Gene* **149**, 237–244 (1994).
16. Bolger, G. B. The PDE4 cAMP-Specific Phosphodiesterases: Targets for Drugs with Antidepressant and Memory-Enhancing Action. *Adv. Neurobiol.* **17**, 385–409 (2017).
17. Richter, W. & Conti, M. Dimerization of the type 4 cAMP-specific phosphodiesterases is mediated by the upstream conserved regions (UCRs). *J. Biol. Chem.* **277**, 40212–40221 (2002).
18. Richter, W. & Conti, M. The oligomerization state determines regulatory properties and inhibitor sensitivity of type 4 cAMP-specific phosphodiesterases. *J. Biol. Chem.* **279**, 30338–30348 (2004).
19. Bolger, G. B. *et al.* Dimerization of cAMP phosphodiesterase-4 (PDE4) in living cells requires interfaces located in both the UCR1 and catalytic unit domains. *Cell. Signal.* **27**, 756–769 (2015).
20. Lee, M. E., Markowitz, J., Lee, J. O. & Lee, H. Crystal structure of phosphodiesterase 4D and inhibitor complex. *FEBS Lett.* **530**, 53–58 (2002).
21. Cedervall, P., Aulabaugh, A., Geoghegan, K. F., McLellan, T. J. & Pandit, J. Engineered stabilization and structural analysis of the autoinhibited conformation of PDE4. *Proc. Natl.*

- Acad. Sci.* **2015**, 201419906 (2015).
22. Huston, E. *et al.* The human cyclic AMP-specific phosphodiesterase PDE-46 (HSPDE4A4B) expressed in transfected COS7 cells occurs as both particulate and cytosolic species that exhibit distinct kinetics of inhibition by the antidepressant rolipram. *J. Biol. Chem.* **271**, 31334–31344 (1996).
 23. Rocque, W. J. *et al.* Human recombinant phosphodiesterase 4B2B binds (R)-rolipram at a single site with two affinities. *Biochemistry* **36**, 14250–14261 (1997).
 24. Souness, J. E. & Rao, S. Proposal for pharmacologically distinct conformers of PDE4 cyclic AMP phosphodiesterases. *Cell. Signal.* **9**, 227–236 (1997).
 25. Rocque, W. J. *et al.* Detailed characterization of a purified type 4 phosphodiesterase, HSPDE4B2B: Differentiation of high- and low-affinity (R)-rolipram binding. *Protein Expr. Purif.* **9**, 191–202 (1997).
 26. Bruno, O. *et al.* New selective phosphodiesterase 4D inhibitors differently acting on long, short, and supershort isoforms. *J. Med. Chem.* **52**, 6546–6557 (2009).
 27. Omori, K. & Kotera, J. Overview of PDEs and their regulation. *Circ. Res.* **100**, 309–327 (2007).
 28. Mika, D. & Conti, M. PDE4D phosphorylation: A coincidence detector integrating multiple signaling pathways. *Cell. Signal.* **28**, 719–724 (2016).
 29. MacKenzie, S. J. *et al.* Long PDE4 cAMP specific phosphodiesterases are activated by protein kinase A-mediated phosphorylation of a single serine residue in Upstream Conserved Region 1 (UCR1). *Br. J. Pharmacol.* **136**, 421–433 (2002).
 30. Sette, C., Conti, M. & Chem, M. J. B. Phosphorylation and Activation of a cAMP-specific Phosphodiesterase by the cAMP-dependent Protein Kinase. *J. Biol. Chem.* **271**, 16526–16534 (1996).
 31. Burgin, A. B. *et al.* Design of phosphodiesterase 4D (PDE4D) allosteric modulators for enhancing cognition with improved safety. *Nat. Biotechnol.* **28**, 63–70 (2010).
 32. Carlisle Michel, J. J. *et al.* PKA-phosphorylation of PDE4D3 facilitates recruitment of the

- mAKAP signalling complex. *Biochem. J.* **381**, 587–592 (2004).
33. Byrne, A. M., Elliott, C., Hoffmann, R. & Baillie, G. S. The activity of cAMP-phosphodiesterase 4D7 (PDE4D7) is regulated by protein kinase A-dependent phosphorylation within its unique N-terminus. *FEBS Lett.* **589**, 750–755 (2015).
 34. Dodge, K. L. *et al.* mAKAP assembles a protein kinase A/PDE4 phosphodiesterase cAMP signaling module. *EMBO J.* **20**, 1921–1930 (2001).
 35. Hoffmann, R., Baillie, G. S., MacKenzie, S. J., Yarwood, S. J. & Houslay, M. D. The MAP kinase ERK2 inhibits the cyclic AMP-specific phosphodiesterase HSPDE4D3 by phosphorylating it at ser579. *EMBO J.* **18**, 893–903 (1999).
 36. Hill, E. V. *et al.* Oxidative stress employs phosphatidylinositol 3-kinase and ERK signalling pathways to activate cAMP phosphodiesterase-4D3 (PDE4D3) through multi-site phosphorylation at Ser239 and Ser579. *Cell. Signal.* **18**, 2056–2069 (2006).
 37. Mika, D., Richter, W. & Conti, M. A CaMKII/PDE4D negative feedback regulates cAMP signaling. *Proc. Natl. Acad. Sci.* **112**, 2023–2028 (2015).
 38. Plattner, F. *et al.* The role of ventral striatal cAMP signaling in stress-induced behaviors. *Nat. Rev. Neurosci.* **18**, 1094–1100 (2016).
 39. Fang, R. *et al.* PDK1/Akt/PDE4D axis identified as a target for asthma remedy synergistic with β 2AR agonists by a natural agent arctigenin. *Allergy Eur. J. Allergy Clin. Immunol.* **70**, 1622–1632 (2015).
 40. Kim, M. J. *et al.* Salt-inducible kinase 1 terminates cAMP signaling by an evolutionarily conserved negative-feedback loop in β -cells. *Diabetes* **64**, 3189–3202 (2015).
 41. Klussmann, E. Protein-protein interactions of PDE4 family members - Functions, interactions and therapeutic value. *Cell. Signal.* **28**, 713–718 (2016).
 42. Taskén, K. A. *et al.* Phosphodiesterase 4D and Protein Kinase A Type II Constitute a Signaling Unit in the Centrosomal Area. *J. Biol. Chem.* **276**, 21999–22002 (2001).
 43. Stefan, E. *et al.* Compartmentalization of cAMP-Dependent Signaling by Phosphodiesterase-4D Is Involved in the Regulation of Vasopressin-Mediated Water

- Reabsorption in Renal Principal Cells. *J. Am. Soc. Nephrol.* **18**, 199–212 (2007).
44. Sachs, B. D. *et al.* p75 neurotrophin receptor regulates tissue fibrosis through inhibition of plasminogen activation via a PDE4/cAMP/PKA pathway. *J. Cell Biol.* **177**, 1119–1132 (2007).
 45. Perry, S. J., Baille, G. S., Kohout, T. A., Mcphee, I. & Magiera, M. M. Targeting of Cyclic AMP Degradation to beta 2-Adrenergic Receptors by beta -Arrestins. *Science (80-.)*. **298**, 834–836 (2002).
 46. Bolger, G. B. *et al.* Scanning peptide array analyses identify overlapping binding sites for the signalling scaffold proteins, β -arrestin and RACK1, in cAMP-specific phosphodiesterase PDE4D5. *Biochem. J.* **398**, 23–36 (2006).
 47. Card, G. L. *et al.* Structural basis for the activity of drugs that inhibit phosphodiesterases. *Structure* **12**, 2233–2247 (2004).
 48. Fox, D., Burgin, A. B. & Gurney, M. E. Structural basis for the design of selective phosphodiesterase 4B inhibitors. *Cell. Signal.* **26**, 657–663 (2014).
 49. Gurney, M. E., D’Amato, E. C. & Burgin, A. B. Phosphodiesterase-4 (PDE4) Molecular Pharmacology and Alzheimer’s Disease. *Neurotherapeutics* **12**, 49–56 (2014).
 50. Zhang, C., Xu, Y., Zhang, H.-T., Gurney, M. E. & O’Donnell, J. M. Comparison of the Pharmacological Profiles of Selective PDE4B and PDE4D Inhibitors in the Central Nervous System. *Sci. Rep.* **7**, 40115 (2017).
 51. Gurney, M. E., Cogram, P., Deacon, R. M., Rex, C. & Tranfaglia, M. Multiple Behavior Phenotypes of the Fragile-X Syndrome Mouse Model Respond to Chronic Inhibition of Phosphodiesterase-4D (PDE4D). *Sci. Rep.* **7**, 1–11 (2017).
 52. Higgs, G. Is PDE4 too difficult a drug target? *Curr Opin Investig Drugs* **11**, 495–498 (2010).
 53. Houslay, M. D., Schafer, P., Zhang, K. Y. J. & Houslay, M. D. Phosphodiesterase-4 as a therapeutic target. *Drug Discov. Today* **10**, 1503–1519 (2005).
 54. Ricciarelli, R. & Fedele, E. Phosphodiesterase 4D: An enzyme to remember. *Br. J. Pharmacol.* **172**, 4785–4789 (2015).

55. Heckman, P. R. A., Blokland, A., Bollen, E. P. P. & Prickaerts, J. Phosphodiesterase inhibition and modulation of corticostriatal and hippocampal circuits: Clinical overview and translational considerations. *Neurosci. Biobehav. Rev.* **87**, 233–254 (2018).
56. Carpenter, D. O., Briggs, D. B., Knox, A. P. & Strominger, N. Excitation of area postrema neurons by transmitters, peptides, and cyclic nucleotides. *J Neurophysiol* **59**, 358–369 (1998).
57. Spina, D. PDE4 inhibitors: Current status. *Br. J. Pharmacol.* **155**, 308–315 (2008).
58. Perez-Torres, S. *et al.* Phosphodiesterase type 4 isozymes expression in human brain examined by in situ hybridization histochemistry and[3H]rolipram binding autoradiography. Comparison with monkey and rat brain. *J Chem Neuroanat* **20**, 349–374 (2000).
59. Robichaud, A. *et al.* Assessing the emetic potential of PDE4 inhibitors in rats. *Br. J. Pharmacol.* **135**, 113–118 (2002).
60. Robichaud, A. *et al.* Deletion of phosphodiesterase 4D in mice shortens alpha(2)-adrenoceptor-mediated anesthesia, a behavioral correlate of emesis. *J. Clin. Investig.* **110**, 1045–1052
61. García-Osta, A., Cuadrado-Tejedor, M., García-Barroso, C., Oyarzábal, J. & Franco, R. Phosphodiesterases as therapeutic targets for Alzheimer’s disease. *ACS Chem. Neurosci.* **3**, 832–844 (2012).
62. Barnes, P. . How corticosteroids control inflammation. *Br. J. Pharmacol.* **148**, 245–254 (2006).
63. Ariga, M. *et al.* Nonredundant function of phosphodiesterases 4D and 4B in neutrophil recruitment to the site of inflammation. *J. Immunol.* **173**, 7531–7538 (2004).
64. Calverley, P. M. A. *et al.* Effect of 1-year treatment with roflumilast in severe chronic obstructive pulmonary disease. *Am. J. Respir. Crit. Care Med.* **176**, 154–161 (2007).
65. Bernabeu, R. *et al.* Involvement of hippocampal cAMP/cAMP-dependent protein kinase signaling pathways in a late memory consolidation phase of aversively motivated learning in rats. *Proc. Natl. Acad. Sci. U. S. A.* **94**, 7041–7046 (1997).

66. Abel, T. *et al.* Genetic demonstration of a role for PKA in the late phase of LTP and in hippocampus-based long-term memory. *Cell* **88**, 615–626 (1997).
67. Barad, M., Bourtchouladze, R., Winder, D. G., Golan, H. & Kandel, E. Rolipram, a type IV-specific phosphodiesterase inhibitor, facilitates the establishment of long-lasting long-term potentiation and improves memory. *Proc. Natl. Acad. Sci. U. S. A.* **95**, 15020–15025 (1998).
68. Rutten, K., Prickaerts, J. & Blokland, A. Rolipram reverses scopolamine-induced and time-dependent memory deficits in object recognition by different mechanisms of action. *Neurobiol. Learn. Mem.* **85**, 132–138 (2006).
69. Bourtchouladze, R. *et al.* A mouse model of Rubinstein-Taybi syndrome: defective long-term memory is ameliorated by inhibitors of phosphodiesterase 4. *Proc. Natl. Acad. Sci. U. S. A.* **100**, 10518–10522 (2003).
70. Lynch, D. C. *et al.* Identification of Novel Mutations Confirms Pde4d as a Major Gene Causing Acrodysostosis. *Hum. Mutat.* **34**, 97–102 (2013).
71. Silve, C. Acrodysostosis: A new form of pseudohypoparathyroidism? *Ann. Endocrinol. (Paris)*. **76**, 110–112 (2015).
72. Bruno, O. *et al.* GEBR-7b, a novel PDE4D selective inhibitor that improves memory in rodents at non-emetic doses. *Br. J. Pharmacol.* **164**, 2054–2063 (2011).
73. Prickaerts, J., Heckman, P. R. A. & Blokland, A. Investigational phosphodiesterase inhibitors in phase I and phase II clinical trials for Alzheimer’s disease. *Expert Opin. Investig. Drugs* **26**, 1033–1048 (2017).
74. Darnell, J. C. *et al.* FMRP stalls ribosomal translocation on mRNAs linked to synaptic function and autism. *Cell* **146**, 247–261 (2011).
75. Lagerbauer, B., Ostareck, D., Keidel, E. M., Ostareck-Lederer, A. & Fischer, U. Evidence that fragile X mental retardation protein is a negative regulator of translation. *Hum. Mol. Genet.* **10**, 329–338 (2001).
76. Berry-Kravis, E., Hicar, M. & Ciurlionis, R. Reduced cyclic AMP production in fragile X syndrome: cytogenetic and molecular correlations. *Pediatr. Res.* **38**, 638–643 (1995).

77. Zhang, H. T. *et al.* Antidepressant-like effects of PDE4 inhibitors mediated by the high-affinity rolipram binding state (HARBS) of the phosphodiesterase-4 enzyme (PDE4) in rats. *Psychopharmacology (Berl)*. **186**, 209–217 (2006).
78. Bruno, O. *et al.* Synthesis and biological evaluation of neutrophilic inflammation inhibitors. *Farmaco* **59**, 223–235 (2004).
79. Brullo, C. *et al.* Synthesis, Biological Evaluation, and Molecular Modeling of New 3-(Cyclopentyloxy)-4-methoxybenzaldehyde O-(2-(2,6-Dimethylmorpholino)-2-oxoethyl) Oxime (GEBR-7b) Related Phosphodiesterase 4D (PDE4D) Inhibitors. *J. Med. Chem.* **57**, 7061–7072 (2014).
80. Brullo, C. *et al.* New insights into selective PDE4D inhibitors : 3- (Cyclopentyloxy) -4-oxime (GEBR-7b) structural development and promising activities to restore memory impairment. *Eur. J. Med. Chem.* **124**, 82–102 (2016).
81. Brullo, C. *et al.* Synthesis, biological activities and pharmacokinetic properties of new fluorinated derivatives of selective PDE4D inhibitors. *Bioorg. Med. Chem.* **23**, 3426–3435 (2015).
82. Ricciarelli, R. *et al.* Memory-enhancing effects of GEBR-32a, a new PDE4D inhibitor holding promise for the treatment of Alzheimer's disease. *Sci. Rep.* **7**, 46320 (2017).
83. Hsiao, K. *et al.* Correlative memory deficits, Ab elevation and amyloid plaques in transgenic mice. *Science (80-.)*. **274**, 99–102 (1996).
84. Prosdocimi, T. *et al.* Molecular Bases of PDE4D Inhibition by Memory-Enhancing GEBR Library Compounds. *Biochemistry* **57**, 2876–2888 (2018).
85. Slack, J. M., Kuzio, J. & Faulkner, P. Characterization of v-cath, a cathepsin L-like proteinase expressed by the baculovirus *Autographa californica* multiple nuclear polyhedrosis virus. *J. Gen. Virol.* **76**, 1091–1098 (1995).
86. Hom, L. G. & Volkman, L. E. *Autographa californica* M nucleopolyhedrovirus chiA is required for processing of V-CATH. *Virology* **277**, 178–183 (2000).
87. Hassell, A. M. *et al.* Crystallization of protein-ligand complexes. *Acta Crystallogr. Sect. D*

- Biol. Crystallogr.* **63**, 72–79 (2006).
88. Winn, M. D. *et al.* Overview of the CCP4 suite and current developments. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **67**, 235–242 (2011).
 89. Adams, P. D. *et al.* PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 213–221 (2010).
 90. Hanwell, M. D. *et al.* Avogadro: An advanced semantic chemical editor, visualization, and analysis platform. *J. Cheminform.* **4**, 1–17 (2012).
 91. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D Biol. Crystallogr.* **66**, 486–501 (2010).
 92. Pettersen, E. F. *et al.* UCSF Chimera - A visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
 93. Schrödinger. The PyMOL Molecular Graphics System.
 94. Chock, S. P. & Huang, C. Y. An optimized continuous assay for cAMP phosphodiesterase and calmodulin. *Anal. Biochem.* **138**, 34–43 (1984).
 95. Biasini, M. *et al.* SWISS-MODEL: Modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, 252–258 (2014).
 96. Frisch, M.J., *et al.* Gaussian 09. (2009).
 97. Pronk, S. *et al.* GROMACS 4.5: A high-throughput and highly parallel open source molecular simulation toolkit. *Bioinformatics* **29**, 845–854 (2013).
 98. Decherchi, S., Bottegoni, G., Spitaleri, A., Rocchia, W. & Cavalli, A. BiKi Life Sciences: a New Suite for Molecular Dynamics and Related Methods in Drug Discovery. *J.Chem.Inf.Model.* **58**, 219–224 (2018).
 99. Piana, S., Lindorff-Larsen, K. & Shaw, D. E. How robust are protein folding simulations with respect to force field parameterization? *Biophys. J.* **100**, L47–L49 (2011).
 100. Carzaniga, L. *et al.* Discovery and Optimization of Thiazolidinyl and Pyrrolidinyl Derivatives as Inhaled PDE4 Inhibitors for Respiratory Diseases. *J. Med. Chem.* **60**, 10026–

- 10046 (2017).
101. Wang, H., Robinson, H. & Ke, H. The Molecular Basis for Different Recognition of Substrates by Phosphodiesterase Families 4 and 10. *J. Mol. Biol.* **371**, 302–307 (2007).
 102. Müller, K., Faeh, C. & Diederich, F. Fluorine in pharmaceuticals: looking beyond intuition. *Science* **317**, 1881–6 (2007).
 103. Huang, A. *et al.* Regioselective Synthesis, NMR, and Crystallographic Analysis of N1-Substituted Pyrazoles. *J. Org. Chem.* **82**, 8864–8872 (2017).
 104. Hatzelmann, A. *et al.* The preclinical pharmacology of roflumilast - A selective, oral phosphodiesterase 4 inhibitor in development for chronic obstructive pulmonary disease. *Pulm. Pharmacol. Ther.* **23**, 235–256 (2010).
 105. Roberts, R. S. *et al.* 4-Amino-7,8-dihydro-1,6-naphthyridin-5(6H)-ones as inhaled phosphodiesterase type 4 (PDE4) inhibitors: structural biology and structure-activity relationships (SAR). *J. Med. Chem.* **5**, acs.jmedchem.7b01751 (2018).
 106. Leckband, D. E. & de Rooij, J. Cadherin Adhesion and Mechanotransduction. *Annu. Rev. Cell Dev. Biol.* **30**, 291–315 (2014).
 107. Shapiro, L. & Weis, W. I. Structure and biochemistry of cadherins and catenins. *Cold Spring Harb. Perspect. Biol.* **1**, (2009).
 108. Parisini, E., Higgins, J. M. G., Liu, J. huan, Brenner, M. B. & Wang, J. huai. The Crystal Structure of Human E-cadherin Domains 1 and 2, and Comparison with other Cadherins in the Context of Adhesion Mechanism. *J. Mol. Biol.* **373**, 401–411 (2007).
 109. Kudo, S., Caaveiro, J. M. M. & Tsumoto, K. Adhesive Dimerization of Human P-Cadherin Catalyzed by a Chaperone-like Mechanism. *Structure* **24**, 1523–1536 (2016).
 110. Dalle Vedove, A., Lucarelli, A. P., Nardone, V., Martino, A. & Parisini, E. The X-ray structure of human P-cadherin EC1-EC2 in a closed conformation provides insight into the type i cadherin dimerization pathway. *Acta Crystallogr. Sect. FStructural Biol. Commun.* **71**, 371–380 (2015).
 111. Brasch, J., Harrison, O. J., Honig, B. & Shapiro, L. Thinking outside the cell: how chdgerins

- drive adhesion. *Trends Cell Biol.* **22**, 299–310 (2012).
112. Harrison, O. J. *et al.* The extracellular architecture of adherens junctions revealed by crystal structures of type I cadherins. *Structure* **19**, 244–256 (2011).
113. Berx, G. & van Roy, F. Involvement of members of the cadherin superfamily in cancer. *Cold Spring Harb. Perspect. Biol.* **1**, (2009).
114. Augustine, C. K. *et al.* Targeting N-cadherin enhances antitumor activity of cytotoxic therapies in melanoma treatment. *Cancer Res.* **68**, 3777–3784 (2008).
115. Doro, F. *et al.* Computational design of novel peptidomimetic inhibitors of cadherin homophilic interactions. *Org. Biomol. Chem.* **13**, 2570–2573 (2015).
116. Nardone, V. *et al.* Crystal Structure of Human E-Cadherin-EC1EC2 in Complex with a Peptidomimetic Competitive Inhibitor of Cadherin Homophilic Interaction. *J. Med. Chem.* **59**, 5089–5094 (2016).
117. Foadi, J. *et al.* Clustering procedures for the optimal selection of data sets from multiple crystals in macromolecular crystallography. *Acta Crystallogr. Sect. D* **69**, 1617–1632 (2013).

Acknowledgments

My PhD scholarship has been funded by the Italian Institute of Technology. All my activities have been conducted at the Center for Nano Science and Technology (IIT) @ PoliMi, in the laboratory and under the supervision of Dr. Emilio Parisini.

The project has also been developed thanks to a fruitful collaboration with:

- Prof. Olga Bruno and Prof. Chiara Brullo, from the University of Genova, who own the property of the GEBR-library and that allowed us to use their compounds for our studies.
- Dr. Paola Storici and Mrs. Marta Semrau, from the Protein Facility of Elettra Sincrotrone Trieste, who supported us with PDE4D3 expression.
- Prof. Andrea Cavalli and Dr. Luca Mollica, from IIT Genova, who carried out the molecular dynamic simulations.
- SLS synchrotron @ PSI, which provided us with beamtime for the X-ray crystallography.

I would sincerely thank:

- Dr. Emilio Parisini, for giving me the possibility to work in his lab, for his teachings and for all the numerous scientific discussions that progressively allowed me to explore my ideas and develop my project. I also want to thank him for having patience with my anxieties and for supporting me during the bad moments, technically and humanly.
- Andrea Dalle Vedove and Stefano Donini for all their lab teachings, tutoring and trainings.
- Andrea Dalle Vedove, Chiara Vettraino, Archimede Torretta and Sara Abbate: for all the fun in lab, all the complaints about experiments not working and for the shared struggling until the result.
- My family that supported me in all my studies

My most precious “thank” goes to Chiara. Her fantastic presence and her constant support has been, and will always be, my driving-force that allow me to be focused on my work and to face all the difficulties.